

**A TIME SERIES ANALYSIS OF ROAD TRAFFIC
FATALITIES IN MALAYSIA**

YUSRIA DARMA

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2017

**A TIME SERIES ANALYSIS OF ROAD TRAFFIC
FATALITIES IN MALAYSIA**

YUSRIA DARMA

**THESIS SUBMITTED IN FULFILMENT OF THE
REQUIREMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2017

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Yusria Darma

(Passport No:

Matric No: KHA070031

Name of Degree: Ph.D.

Title of Thesis:

A Time Series Analysis of Road Traffic Fatalities in Malaysia

Field of Study: Transportation Engineering

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature

Date:

Subscribed and solemnly declared before,

Witness's Signature

Date:

Name : Ir. Sulaiman Abdullah

Designation : Accredited Road Safety Auditor by JKR Malaysia

ABSTRACT

In Malaysia, with the current road safety developments, it is perceived that the target of the latest Road Safety Plan of Malaysia 2014–2020 that is the main objective is to reduce the deaths in 2020 by 5,358 deaths may not be achieved. Even though various interventions and road safety measures have been implemented through the development of legislations, standards, guidelines and integrated road safety programmes. The primary objectives of this study are: (1) to describe the characteristics of road safety in Malaysia, (2) to investigate the impact of road safety measures in reducing the rate of fatalities, (3) to investigate the factors that influencing the rate of fatalities and (4) to develop time series models to predict the rate of road traffic fatalities in Malaysia. In order to achieve these objectives, three forecasting models are developed based on the time series analysis technique, namely: (1) autoregressive integrated moving average model, (2) transfer function-noise model and (3) state-space model. The multiple regression is used to select the explanatory variables that are correlated significantly with the number of road traffic fatalities. Then, these variables are the input variables for the state-space model. The effectiveness of a road safety measure is checked with the autoregressive integrated moving average and transfer function-noise model. Whilst, for forecasting the fatalities up to year 2020, the autoregressive integrated moving average, transfer function-noise and state-space models are employed. Based on the results, the characteristics of the current road safety in Malaysia are: (1) the major victims of road traffic accidents are motorcyclists, (2) young adult drivers/riders aged 16–25 years make up the highest percentage of the total fatalities, with a value of 35%, and (3) the highest rate of fatal accidents per kilometre occurs at expressways. In addition, the effectiveness of a number of road safety measures is also investigated in this study. The results show that the enactment of the Seat Belt Rules in 1978 decreases the rate of car driver fatalities by 58%. However, the Motorcycle Daytime Running Headlight Regulation in 1992 is not an

effective road safety measure to reduce the rate of motorcycle fatalities throughout the nation. In contrast, the Integrated Road Safety Operations (*Ops Sikap*) is an effective measure. The National Road Safety Plan 2006–2010 decreases the rate of the total road traffic fatalities by only 9%. This achievement is rather low compared to the targeted value of 52.4%. The results show that the significant explanatory variables to forecast the fatalities are: (1) the number of hospital beds per 1,000 people, (2) the percentage of registered motorcycles and (3) road length. Moreover, it is expected that the targeted rate of fatalities of 2.0 per 10,000 registered vehicles will be achieved by year 2023 with intensive enforcement. However, drastic actions need to be taken to achieve the target in the Road Safety Plan 2014–2020.

Keywords: road, safety, fatalities, time series model, Malaysia

ABSTRAK

Di Malaysia, dengan perkembangan keselamatan jalan raya sekarang, sasaran dalam Pelan Keselamatan Jalan Raya 2014–2020 yang mempunyai objektif utama untuk mengurangkan kematian pada tahun 2020 menjadi 5,358 mungkin tidak dapat dicapai. Walaupun pelbagai intervensi dan langkah keselamatan jalan raya telah diusahakan seperti penggubalan undang-undang, pembangunan piawaian dan garis panduan keselamatan jalan raya, serta pelaksanaan program keselamatan jalan raya bersepadu. Tujuan kajian ini dijalankan adalah untuk: (1) menerangkan ciri-ciri keadaan keselamatan jalan raya di Malaysia, (2) menyiasat kesan langkah keselamatan jalan raya dalam mengurangkan kadar kematian, (3) mengkaji faktor-faktor yang mempengaruhi kadar kematian dan (4) membangunkan model siri masa bagi meramal kadar kematian yang disebabkan oleh kemalangan jalan raya. Untuk mencapai objektif di atas, tiga model peramalan telah dibangunkan berdasarkan kaedah analisis siri masa, iaitu: (1) model purata bergerak autoregresif terkamir (2) model fungsi pindah-gangguan dan (3) model keadaan-ruang. Regresi berganda digunakan untuk memilih pemboleh ubah yang berkorelasi dengan ketara dengan jumlah kematian di jalan raya. Kemudian, pemboleh ubah ini menjadi input untuk model keadaan-ruang. Keberkesanan langkah keselamatan jalan raya diperiksa dengan model purata bergerak autoregresif terkamir dan model fungsi pindah-gangguan. Untuk meramalkan kematian hingga tahun 2020, ketiga model adalah digunakan. Hasil kajian menunjukkan bahawa ciri-ciri keadaan keselamatan jalan raya di Malaysia adalah seperti berikut: (1) kebanyakan mangsa kemalangan jalan raya adalah penunggang motosikal, (2) pemandu/penunggang dewasa muda yang berumur antara 16–25 tahun mencatat peratus kematian di jalan raya tertinggi dengan jumlah sebanyak 35%, dan (3) kadar kemalangan maut bagi setiap kilometer adalah paling tinggi di lebuh raya. Keberkesanan beberapa langkah keselamatan jalan raya juga telah disiasat dalam kajian ini. Hasil kajian menunjukkan bahawa enakmen Peraturan Tali Pinggang Keledar pada

tahun 1978 telah berjaya mengurangkan kadar kematian bagi pemandu kereta sebanyak 58%. Walau bagaimanapun, peraturan yang mewajibkan penunggang motosikal memasang lampu pada tahun 1992 didapati tidak berkesan untuk mengurangkan kadar kematian penunggang motosikal di seluruh negara. Namun begitu, Operasi Bersepadu Keselamatan Jalan Raya (Ops Sikap) merupakan langkah yang berkesan. Pelan Keselamatan Jalan Raya Malaysia 2006–2010 dapat mengurangkan kadar kematian di jalan raya sebanyak 9% sahaja. Pencapaian ini tidak begitu memberangsangkan berbanding dengan sasaran, iaitu pengurangan sebanyak 52.4%. Hasil kajian menunjukkan bahawa pemboleh ubah penjelas signifikan untuk meramal bilangan kematian adalah seperti berikut: (1) bilangan katil di hospital bagi setiap 1,000 penduduk, (2) peratus motosikal berdaftar dan juga (3) panjang jalan. Tambahan pula, sasaran kadar kematian 2.0 bagi setiap 10,000 kenderaan berdaftar dijangka akan dicapai pada tahun 2023 dengan penguatkuasaan yang intensif. Namun begitu, tindakan drastik perlu diambil untuk mencapai sasaran dalam Pelan Keselamatan Jalan Raya 2014–2020.

Kata Kunci: jalan raya, keselamatan, kematian, model siri masa, Malaysia

ACKNOWLEDGEMENT

First and foremost, all praise be to Allah the Almighty for granting me with the knowledge, strength, courage, patience and perseverance to complete my doctoral degree safely. It has indeed been one long, arduous journey, which I believe that it would have been impossible without His blessings and mercy.

This study would not have been successful without the contribution of many significant individuals. I shall begin by expressing my utmost gratitude to my supervisor, Prof. Ir. Dr. Mohamed Rehan Karim, for his professionalism, continuing guidance and support which have enabled me to complete my study successfully. My heartfelt thanks goes to both my internal examiner, Dr. Mastura Binti Adam (University of Malaya) and external examiners, Prof. Dr. András Várhelyi (Lund University, Sweden) and Prof. Dr. Mohammed A Quddus (Loughborough University, UK) for their constructive criticisms and comments which have helped me improve the overall quality of this thesis.

Indeed, research would be nearly impossible to carry out (not to mention monotonous) if it is done entirely by myself. For this reason, I wish to express my cordial thanks to my fellow postgraduate mates in the Traffic Laboratory for their invaluable assistance and moral support, making this whole journey fun and bearable. My special thanks goes to the technical and administrative staff of Department of Civil Engineering, University of Malaya, for their assistance and support throughout the course of my study. I am greatly indebted to my family and friends for their endless love, prayers, encouragement and support. Last but not least, may this humble contribution benefits everyone (directly and indirectly) and help reduce the number of road fatalities in Malaysia.

TABLE OF CONTENTS

| | |
|---|------|
| Title Page | i |
| Original Literary Work Declaration | ii |
| Abstract | iii |
| Abstrak | v |
| Acknowledgement | vii |
| Table of Contents | viii |
| List of Figures | xvii |
| List of Tables | xxiv |
| List of Abbreviations | xxvi |
| CHAPTER 1: INTRODUCTION | 1 |
| 1.1 Background of the study | 1 |
| 1.2 Road safety development in Malaysia | 3 |
| 1.3 Problem Statement..... | 6 |
| 1.4 Objectives and significance of the study..... | 10 |
| 1.5 Research questions..... | 11 |
| 1.6 Methodological framework..... | 12 |
| 1.7 Organization of the thesis..... | 16 |
| 1.8 Chapter summary..... | 17 |
| CHAPTER 2: LITERATURE REVIEW | 19 |
| 2.1 Road safety research development in early years..... | 19 |
| 2.2 Seminal works in road safety research and development | 21 |
| 2.3 Explanatory variables appear to affect fatalities | 24 |
| 2.3.1 Socioeconomic and demographic indicators | 26 |
| 2.3.1.1 GDP and household income | 26 |
| 2.3.1.2 Vehicle ownership | 29 |

| | | |
|---------|--|----|
| 2.3.1.3 | Unemployment rate..... | 29 |
| 2.3.1.4 | Economic growth..... | 30 |
| 2.3.1.5 | Fuel price..... | 31 |
| 2.3.2 | Age, gender categories and urbanization..... | 32 |
| 2.3.3 | Weather and time variations | 33 |
| 2.3.4 | Medical care facilities..... | 34 |
| 2.3.5 | Geographical factors and road geometry | 35 |
| 2.3.6 | Proportion of motorcycle and public transit user | 36 |
| 2.3.7 | Other explanatory variables | 37 |
| 2.4 | The effect of fuel shortages on road safety in 1973..... | 38 |
| 2.5 | The decrease in death tolls due to road traffic accidents in 1972–1973 | 39 |
| 2.6 | Road safety measure and intervention | 40 |
| 2.6.2 | Maximum blood alcohol concentration..... | 44 |
| 2.6.3 | Speed limit | 45 |
| 2.6.4 | Speed Limit Enforcement Camera | 46 |
| 2.6.5 | Provision for motorcyclist | 47 |
| 2.6.6 | Driver licensing..... | 49 |
| 2.6.7 | Laws enforcement | 49 |
| 2.7 | Road safety policy and target | 50 |
| 2.7.1 | Sustainable safety vision..... | 51 |
| 2.7.2 | Road safety target..... | 52 |
| 2.7.2.1 | Setting road safety target..... | 52 |
| 2.7.2.2 | Vision zero | 53 |
| 2.8 | Recent road safety reports by WHO | 54 |
| 2.9 | Risk compensation..... | 56 |
| 2.9.1 | Researchers who opposed the risk compensation theory | 56 |

| | | |
|--------|---|-----|
| 2.10 | Unit of exposure to risk..... | 61 |
| 2.11 | Road safety index | 64 |
| 2.12 | Road safety modelling | 65 |
| 2.12.1 | Road safety modelling is complex | 66 |
| 2.12.2 | Road safety modelling is practical | 67 |
| 2.12.3 | Exponential model..... | 69 |
| 2.12.4 | Extrapolation model | 70 |
| 2.12.5 | Structural time series model..... | 71 |
| 2.12.6 | Time series analysis and intervention model | 74 |
| 2.12.7 | State-space model..... | 76 |
| 2.12.8 | Model with stratified data | 77 |
| 2.13 | Key studies which involve the use of time series modelling to analyse and predict road safety | 79 |
| 2.13 | Recent time series models used in road safety | 89 |
| 2.14 | Statistical methods used to forecast time series data | 92 |
| 2.14.1 | Univariate and multivariate methods | 92 |
| 2.14.2 | Deterministic and stochastic models | 93 |
| 2.14.3 | Microscopic and macroscopic levels..... | 94 |
| 2.14.4 | Time series data patterns | 94 |
| 2.14.5 | Models used for analysing and forecasting road traffic casualties | 95 |
| 2.15 | Time series models used in this study | 96 |
| 2.15.1 | Box-Jenkins (ARIMA) models | 97 |
| 2.15.2 | Intervention analysis and transfer function-noise models | 98 |
| 2.15.3 | State-space models | 99 |
| 2.15.4 | Least squares and maximum likelihood methods | 100 |
| 2.16 | Chapter summary..... | 100 |

| | |
|--|------------|
| CHAPTER 3: RESEARCH METHODOLOGY | 102 |
| 3.1 Construction of models | 102 |
| 3.2 Research data and its reliability | 105 |
| 3.3 Data analysis tools | 108 |
| 3.4 Selection of a forecasting method | 108 |
| 3.5 Steps involved in forecasting | 109 |
| 3.6 Descriptive statistical procedures | 111 |
| 3.6.1 Numerical summary | 111 |
| 3.6.2 Probability distributions | 112 |
| 3.6.3 Sampling distribution | 113 |
| 3.6.4 Estimation | 115 |
| 3.6.5 Hypothesis testing | 115 |
| 3.6.6 Correlation analysis | 116 |
| 3.6.7 Scatter diagrams | 116 |
| 3.6.8 Correlation coefficient | 117 |
| 3.7 Examination of data patterns and selection of forecasting methods..... | 118 |
| 3.7.1 Examination of time series data patterns | 118 |
| 3.7.2 Examination of data patterns with autocorrelation analysis | 118 |
| 3.7.3 Forecasting errors..... | 121 |
| 3.8 Components of a time series | 122 |
| 3.8.1 Decomposition | 122 |
| 3.8.2 Forecasting trend..... | 125 |
| 3.8.3 Seasonality | 125 |
| 3.9 Regression analysis..... | 126 |
| 3.9.1 Inference for the regression models | 127 |
| 3.9.2 Multi-collinearity | 128 |

| | | |
|--|--|------------|
| 3.9.3 | Stepwise regression | 129 |
| 3.10 | Regression of time series data | 131 |
| 3.10.1 | The autocorrelation problem | 131 |
| 3.10.2 | Test for serial correlation | 133 |
| 3.10.3 | Solutions for autocorrelation problems | 135 |
| 3.10.3.1 | Box-Cox transformation | 135 |
| 3.10.3.2 | Regression with differencing | 136 |
| 3.10.3.3 | Autoregressive models | 138 |
| 3.10.4 | The heteroscedasticity problem | 139 |
| 3.11 | Box-Jenkins (ARIMA) models | 139 |
| 3.11.1 | Box-Jenkins (ARIMA) methodology | 140 |
| 3.11.1.1 | Autoregressive models | 142 |
| 3.11.1.2 | Moving average models | 143 |
| 3.11.1.3 | Autoregressive moving average models | 145 |
| 3.11.1.4 | Model construction strategy | 146 |
| 3.11.1.5 | Model selection criteria | 153 |
| 3.11.1.6 | Summary of the ARIMA model construction | 154 |
| 3.12 | ARMA(p, q) intervention analysis | 155 |
| 3.12 | Transfer function-noise model | 158 |
| 3.13 | State-space model | 162 |
| 3.14 | Chapter summary | 165 |
| CHAPTER 4: RESULTS AND ANALYSIS | | 167 |
| 4.1 | Key statistics on the road traffic safety scenario in Malaysia | 167 |
| 4.1.1 | Road traffic fatalities according to the type of road users | 169 |
| 4.1.2 | Fatal accidents in urban and rural areas | 171 |
| 4.1.3 | Driver/rider fatalities according to gender | 173 |

| | | |
|-------|---|-----|
| 4.1.4 | Driver/rider fatalities according to age group | 174 |
| 4.1.5 | Fatal accidents according to road category | 176 |
| 4.1.6 | Fatal accidents according to the type of road segment | 178 |
| 4.1.7 | Fatal accidents according to the type of junction | 180 |
| 4.1.8 | Driver/rider fatalities according to traffic system | 181 |
| 4.1.9 | Fatal accidents according to the time of day | 183 |
| 4.2 | Investigation of the government interventions and road safety measures on road safety improvement | 184 |
| 4.2.1 | Legislations on the road traffic safety | 184 |
| 4.2.2 | Standards | 186 |
| 4.2.3 | Guidelines | 186 |
| 4.2.4 | National road safety targets and plans | 186 |
| 4.2.5 | Programmes | 187 |
| 4.2.6 | Formation of government agencies | 187 |
| 4.3 | Modelling the impact of the interventions | 190 |
| 4.3.1 | Data for the univariate time series analysis | 190 |
| 4.3.2 | Data series plots | 192 |
| 4.3.3 | ARIMA modelling | 198 |
| 4.3.4 | Intervention analysis modelling | 210 |
| 4.3.5 | Impact of the road safety measures or interventions | 229 |
| 4.4 | Forecasting the number and rate of fatalities | 234 |
| 4.4.1 | ARIMA model | 235 |
| 4.4.2 | Transfer function-noise model | 236 |
| 4.4.3 | Forecasting using the state-space model | 239 |

| | | |
|--|---|------------|
| 5.2.1.7 | Road Safety Programmes in 1997 | 277 |
| 5.2.1.8 | Integrated Road Safety Operations (Ops Sikap) in 2001 | 279 |
| 5.2.1.9 | Road Safety Plan 2006–2010 | 282 |
| 5.2.2 | Which is the best time series model among the three models developed in this study in order to predict the rate of fatalities in Malaysia?..... | 283 |
| 5.2.3 | Is it acceptable to only use the descriptive/univariate model (without explanatory variables) which is practised extensively in Netherlands? .. | 285 |
| 5.2.4 | What are the significant explanatory variables which affect the number of fatalities?..... | 285 |
| 5.2.5 | Is it possible for the number of fatalities to increase significantly on an ongoing basis considering the fact that most of road safety measures have been implemented?..... | 287 |
| 5.2.6 | Is the major decline of fatalities in 1997–1998 due to the initiatives and measures implemented in Malaysia? Or is it possible that this decline is a consequence of an unexpected regional economic crisis at the time which in turn reduces traffic exposure (vehicle-kilometres travelled or fuel consumption)?..... | 287 |
| 5.2.7 | Is it practical to include only the population and number of vehicles as the explanatory variables as had been done by MIROS?..... | 288 |
| 5.2.8 | Can the number of fatalities targeted in the Road Safety Plan 2014–2020 (death toll: 5,358) be achieved in year 2020? | 289 |
| 5.2.9 | Has the risk compensation theory occurred? | 290 |
| CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS..... | | 291 |
| 6.1 | Conclusions | 291 |
| 6.2 | Recommendations..... | 295 |

| | | |
|-----|---|-----|
| 6.3 | Contribution of the study | 296 |
| 6.4 | Limitation of the study..... | 297 |
| 6.5 | Recommendation for future research..... | 297 |
| | REFERENCES..... | 298 |
| | APPENDICES..... | 325 |

University of Malaya

LIST OF FIGURES

| | | |
|------------|---|-----|
| Figure 1.1 | Number and rate of road fatalities in Malaysia from 1981 to 2012 | 4 |
| Figure 1.2 | Rate of fatalities in countries with a vehicle ownership rate greater than 0.5 in year 2010 | 7 |
| Figure 1.3 | Trend of rate of fatalities in 2010 based on Smeed's and Koren- Borsos's models | 9 |
| Figure 1.4 | Conceptual framework of the methodology adopted in this study | 15 |
| Figure 3.1 | Conceptual framework of the modelling process | 104 |
| Figure 3.2 | Normal distribution curve with population mean μ | 114 |
| Figure 3.3 | Steps involved in the iterative approach to construct a Box- Jenkins model for forecasting or control | 142 |
| Figure 3.4 | Examples of ACFs and PACFs of common ARMA models | 149 |
| Figure 3.5 | Example of a non-stationary data series and its ACF and PACF pattern | 150 |
| Figure 3.6 | Typical intervention functions | 157 |
| Figure 3.7 | Typical response to a step input | 158 |
| Figure 3.8 | Examples of impulse and step response functions | 161 |
| Figure 4.1 | Rate of road traffic fatalities per 10,000 registered vehicles and rate of road traffic fatalities per 100,000 population due to road traffic accidents in Malaysia from 1981 to 2012 | 169 |
| Figure 4.2 | Number of fatalities according to the type of road users from 1983 to 2012 | 170 |

| | | |
|-------------|---|-----|
| Figure 4.3 | Rate of fatalities per 10,000 registered vehicles for selected road users from 1981 to 2012 | 171 |
| Figure 4.4 | Percentage of fatal accidents from 1993 to 2012 | 172 |
| Figure 4.5 | Rate of fatal accidents in urban and rural areas from 1993 to 2012 | 173 |
| Figure 4.6. | Percentage of male and female driver/rider fatalities from 1993 to 2012 | 174 |
| Figure 4.7 | Rate of male and female driver/rider fatalities from 1993 to 2012 | 174 |
| Figure 4.8 | Percentage of driver/rider fatalities from 1993 to 2012 according to age group | 175 |
| Figure 4.9 | Rate of driver/rider fatalities from 1993 to 2012 according to age group | 176 |
| Figure 4.10 | Percentage of fatal accidents from 1993 to 2012 according to road category | 177 |
| Figure 4.11 | Rate of fatal accidents from 1993 to 2012 according to road category | 178 |
| Figure 4.12 | Percentage of fatal accidents from 1993 to 2012 according to the type of road segment | 179 |
| Figure 4.13 | Rate of fatal accidents from 1993 to 2012 according to the type of road segment | 179 |
| Figure 4.14 | Percentage of fatal accidents from 1993 to 2012 according to the type of junction | 180 |
| Figure 4.15 | Rate of fatal accidents from 1993 to 2012 according to the type of junction | 181 |

| | | |
|-------------|--|-----|
| Figure 4.16 | Percentage of driver/rider fatalities from 1993 to 2012 according to traffic system | 182 |
| Figure 4.17 | Rate of driver/rider fatalities from 1993 to 2012 according to traffic system | 182 |
| Figure 4.18 | Percentage of fatal accidents from 1993 to 2012 according to the time of day | 183 |
| Figure 4.19 | Rate of fatal accidents from 1993 to 2012 according to the time of day | 184 |
| Figure 4.20 | Rate of motorcyclist fatalities from 1969 to 1988 and the year in which the safety helmet rules came into force | 193 |
| Figure 4.21 | Rate of motorcar driver fatalities from 1969 to 1988 and the year in which the seat belt rules came into force | 193 |
| Figure 4.22 | Rate of vulnerable road user fatalities from 1969 to 1988 and the year in which the seat belt rules came into force | 194 |
| Figure 4.23 | Rate of motor vehicle accidents from 1969 to 1994 and the year in which the speed limit rules came into force | 194 |
| Figure 4.24 | Rate of accidents related to alcohol from 1969 to 1996 and the year in which the rules for driving while under the influence of intoxicating liquor or drugs came into force | 195 |
| Figure 4.25 | Rate of motorcycle accidents from 1969 to 1996 and the year in which the regulation for motorcyclists to switch on headlights during daytime was made compulsory | 195 |
| Figure 4.26 | Rate of motorcycle fatalities related to head injuries from 1992 to 2012 and the year in which the second revision of the helmet standards was imposed | 196 |

| | | |
|-------------|---|-----|
| Figure 4.27 | Rate of road traffic fatalities from 1981 to 2000 and the year in which comprehensive safety programmes were implemented ... | 196 |
| Figure 4.28 | Rate of road traffic fatalities from 1981 to 2005 and the year in which the Integrated Road Safety Operations (<i>Ops Sikap</i>) was imposed | 197 |
| Figure 4.29 | Rate of road traffic fatalities from 1981 to 2012 and the year in which the National Road Safety Plan 2006 was launched | 197 |
| Figure 4.30 | SAS syntax used to check stationarity of the rate of fatalities series from 1981 to 2000 | 200 |
| Figure 4.31 | Sample ACF, IACF, PACF and autocorrelation check for white noise of the log rate of fatalities series from 1981 to 2000 | 201 |
| Figure 4.32 | Lambda (λ) value of the log rate of fatalities from 1981 to 2000 | 202 |
| Figure 4.33 | Results of the autocorrelation check for white noise for the log rate of fatalities from 1981 to 2000 | 202 |
| Figure 4.34 | SAS syntax used to conduct the Augmented Dickey-Fuller (ADF) test for the log rate of fatalities from 1981 to 2000 | 203 |
| Figure 4.35 | Results of the ADF test for log rate of fatalities without differencing | 203 |
| Figure 4.36 | Results of the ADF test and autocorrelation check for white noise for the log rate of fatalities series with first-order differencing | 204 |
| Figure 4.37 | SAS programming syntax used to determine the potential ARIMA(p,d,q) models which will forecast the log rate of fatalities from 1981 to 2000 | 209 |
| Figure 4.38 | SAS syntax used to check the adequacy of the ARIMA(12,1,4) model adequacy forecast the fatalities | 207 |

| | | |
|-------------|--|-----|
| Figure 4.39 | Parameters, residuals and adequacy check results of the ARIMA(12,1,4) model | 208 |
| Figure 4.40 | Forecasted log rate of fatalities produced by the ARIMA(12,1,4) model | 209 |
| Figure 4.41 | Time plot of actual and forecasted log rate of fatalities produced by the ARIMA(12,1,4) model | 209 |
| Figure 4.42 | SAS syntax used to check stationarity in the rate of fatalities from 1981 to 2005 | 210 |
| Figure 4.43 | Sample ACF, IACF, PACF and autocorrelation check for white noise in rate of fatalities from 1981 to 2005 | 211 |
| Figure 4.44 | Lambda (λ) value of the rate of fatalities data from 1981 to 2005 | 213 |
| Figure 4.45 | Autocorrelation check on the log rate of fatalities for white noise | 214 |
| Figure 4.46 | SAS syntax for the Augmented Dickey-Fuller (ADF) test of log rate of fatalities | 214 |
| Figure 4.47 | Results of the Augmented Dickey-Fuller (ADF) test of no differenced series | 215 |
| Figure 4.48 | 4.48 Results of the Augmented Dickey-Fuller (ADF) test with first-order differencing | 215 |
| Figure 4.49 | SAS syntax used to determine the cross-correlation pattern of the model's | 218 |
| Figure 4.50 | Cross-correlation pattern between the variables of the model .. | 218 |
| Figure 4.51 | SAS programming syntax used to develop the potential transfer function-noise model | 220 |

| | | |
|-------------|---|-----|
| Figure 4.52 | Estimated parameters, information criterion and autocorrelation check of residuals for the transfer function-noise (0,1,8) and (3,1,4) models | 223 |
| Figure 4.53 | SAS syntax used to determine the ACF and PACF of the residuals, and the cross-correlation between the residuals and forecasts | 224 |
| Figure 4.54 | Sample of ACF and PACF residual plot for the transfer function-noise (12,1,4) model | 225 |
| Figure 4.55 | Autocorrelation check of residuals and cross-correlation check of residuals for the transfer function-noise (12,1,4) model | 225 |
| Figure 4.56 | Parameters of the transfer function-noise (12,1,4) model | 226 |
| Figure 4.57 | Parameters of the transfer function-noise (12,1,4) model without the constant | 227 |
| Figure 4.58 | Actual and forecasted log rate of fatalities produced by the transfer function-noise (12,1,4) model | 228 |
| Figure 4.59 | Time plot of actual and forecasted log rate of fatalities produced by the transfer function-noise (12,1,4) model | 228 |
| Figure 4.60 | Actual and forecasted rate of fatalities from 1981 to 2020 produced by the ARIMA(2,1,0) model | 236 |
| Figure 4.61 | Actual and forecasted rate of fatalities from 1981 to 2020 produced by the transfer function-noise (5,1,7) model | 238 |
| Figure 4.62 | Actual and forecasted rate of fatalities up to year 2020 produced by the ARIMA(2,1,0) and transfer function-noise (5,1,7) models | 239 |

| | | |
|-------------|--|-----|
| Figure 4.63 | Preliminary estimates for the state-space model | 245 |
| Figure 4.64 | State-space model and fitted model | 246 |
| Figure 4.65 | Actual and forecasted number of road traffic fatalities from 1982 to 2020 produced by the state-space model | 247 |
| Figure 4.66 | Actual and forecasted rate of road traffic fatalities from 1982 to 2020 produced by the state-space model | 247 |

University of Malaya

LIST OF TABLES

| | | |
|------------|---|-----|
| Table 1.1 | List of thesis chapters and the contents of each chapter | 16 |
| Table 2.1 | Key studies on road safety research which involve the use of time series modelling methods | 80 |
| Table 3.1 | Data and its source | 105 |
| Table 3.2 | Box-Cox transformation based on the value of λ | 135 |
| Table 3.3 | Properties of autoregressive (AR), moving average (MA) and mixed autoregressive moving average (ARMA) processes | 146 |
| Table 4.1 | Rate of fatal accidents per kilometre of road in 2012 | 177 |
| Table 4.2 | Type of data used in univariate analysis and its corresponding period of observations | 191 |
| Table 4.3 | Table 4.3 Rate of road traffic fatalities from 1981 to 2000 | 199 |
| Table 4.4 | Rate of road traffic log fatalities from 1981 to 2000 | 200 |
| Table 4.5 | ARIMA models in which the parameters' p -values are less than 0.1 and their information criteria | 206 |
| Table 4.6 | Rate of road traffic log fatalities from 1981 to 2005 | 213 |
| Table 4.7 | Transfer function-noise models with significant parameters ... | 221 |
| Table 4.8 | SBC of the transfer function-noise models | 222 |
| Table 4.9 | Models used to assess the impact of road safety measures | 230 |
| Table 4.10 | Initial and long-term effect of the interventions on road safety .. | 233 |
| Table 4.11 | Details of the transfer function-noise (5,1,7) model | 237 |
| Table 4.12 | VIF and other statistical parameters of explanatory variables .. | 242 |

| | | |
|------------|--|-----|
| Table 4.13 | Variables with p -value less than 0.1 | 243 |
| Table 4.14 | VIF and other statistical parameters of model with six variables | 244 |
| Table 4.15 | VIF and other statistical parameters of model with three variables | 244 |
| Table 5.1 | Forecasted number of fatalities in 2015 and 2020 produced by the three different models developed in this study as well as MIROS | 289 |

LIST OF ABBREVIATIONS

| Abbreviation | Description |
|--------------|---|
| AADT | Annual average daily traffic |
| ABS | Anti-lock brakes |
| ACF | Autocorrelation function |
| ADF | Augmented Dickey-Fuller |
| AES | Automated enforcement system |
| AIC | Akaike information criterion |
| APE | Absolute percentage error |
| ARIMA | Autoregressive integrated moving average |
| ARIMAX | Autoregressive integrated moving average with explanatory variables |
| ASEAN | Association of Southeast Asian Nations |
| ATJ | Arahan teknik jalan |
| BAC | Blood alcohol concentration |
| BIC | Bayesian Information Criterion |
| CCF | Cross-correlations function |
| CDC | Centers for Disease Control and Prevention of the United States |
| CONASET | Comisión Nacional de Seguridad de Tránsito, Chile |
| DEA | Data envelopment analysis |
| DRAG | Demand routière, les accidents et leur gravité |
| DWI | Driving while intoxicated |
| EC | European Commission |
| ECMT | European Conference of Ministers of Transport |

| | |
|-------|---|
| EU | European Union |
| GDP | Gross domestic product |
| GNP | Gross national product |
| IACF | Inverse autocorrelation function |
| ISO | International Standards Organization |
| JKJR | Jabatan Keselamatan Jalan Raya, Malaysia |
| JKR | Jabatan Kerja Raya, Malaysia |
| JPJ | Jabatan Pengangkutan Jalan, Malaysia |
| LRT | Light rail transit |
| MAAP | Microcomputer accident analysis package |
| MAE | Mean absolute error |
| MAPE | Mean average percentage error |
| MASE | Mean absolute scaled error |
| MIROS | Malaysian institute of road safety research |
| MoT | Ministry of Transport, Malaysia |
| MRT | Mass rapid transit |
| MS | Malaysian standard |
| MSP | Motorcycle safety program |
| NARX | Non-linear auto-regression exogenous |
| NGOs | Non-governmental organisations |
| NHTSA | National Highway Traffic Safety Administration |
| NTI | National Transport Insurance, Australia |
| OECD | Organisation for Economic Co-operation and Development |
| ONISR | Observatoire National Interministériel de Sécurité Routière |
| PACF | Partial autocorrelation function |
| PDRM | Polis Diraja Malaysia |

| | |
|----------|---|
| RAE | Relative absolute error |
| REAM | Road Engineering Association of Malaysia |
| RSDI | Road safety development index |
| SBC | Schwarz's Bayesian criterion |
| SPIs | Safety performance indicators |
| SUVs | Sports utility vehicles |
| SWOV | Stichting wetenschappelijk onderzoek verkeersveiligheid, The Netherlands |
| TAG | Transports accidents gravité |
| TRULS | TRafikk, ulykker og skadegrad |
| US | The United States |
| UK | The United Kingdom |
| UN ESCAP | United Nations Economic and Social Commission for Asia and the Pacific |
| US DOT | The United States Department of Transportation |
| VAR | Vector autoregressive |
| VIF | Variance inflation factor |
| VMT | Vehicle mile travelled |
| WHO | World Health Organization |

CHAPTER 1: INTRODUCTION

The chapter begins with a brief background of this study as well as a concise discussion on the road traffic scenario in Malaysia that forms the motivation of this study. The issues that need to be addressed are highlighted and the research objectives are presented. This is followed by the methodological framework used to conduct the study. The organization of this thesis is presented at the end of this chapter.

1.1 Background of the study

At present, there are several differences between the road traffic fatalities in Malaysia and those in developed countries. The absolute number of fatalities in most developed countries reached a peak in between 1970 and 1972 and has declined ever since. However, the number of deaths on roads in Malaysia still increases to date. The victims of road accidents are typically motorcyclists, accounting for up to 60% of the total number of deaths in 2012. The blood alcohol concentration (BAC) may not be a considerable causation factor of fatal road accidents. Only a small number of vehicle occupants (1.3%) have been positively tested to be influenced by alcohol when involved in fatal accidents.

Indeed, various road safety measures that are practised in developed countries which are also implemented in Malaysia. These include enforcement of legislation, education and training, improvements in road safety engineering and media campaigns. Among the major road safety measures implemented are the enactment of motorcycle safety helmet rules in 1973, the enforcement of seat belt use for front seat occupants in 1979, black spots treatment programme since 1995, mandated speeding offences camera, new helmet standards launched in 1996, exclusive lanes for motorcyclists and wide coverage media campaigns since 1997. Road safety auditing has also been imposed for federal route projects and state routes since 1998 and 2007, respectively.

Road traffic accidents have occurred many centuries ago, ever since people travelled with horse and carriages. However, the number of road traffic accidents is expected to escalate rapidly beginning from last decade, especially in low-income and middle-income countries. In 1869, the first known fatal motor vehicle accident occurred in Ireland. Indeed, the accident occurred during a motorcar experiment, whereby the victim was thrown from the seat of the experimental steam-powered car. The victim was crushed by one of the car's heavy wheels (Anonymous, 2007). Following the first road fatality in London in 1896, coroner William Morris said, 'This should never happen again' (Chalmers, 2010). Unfortunately, the coroner's advice shall never become a reality since the number of fatalities continue to escalate rapidly over the years. In fact, road traffic accidents have been listed as one of the ten leading causes of death in the US since 1926 (CDC, 2009) and have achieved the 9th spot as one of the causes of deaths throughout the globe in 2004 (WHO, 2008).

Approximately 1.24 million people die every year from road traffic accidents in 2010, whereas another 20 to 50 million people sustain non-fatal injuries resulting from vehicle collisions (WHO, 2013). The number of road traffic deaths annually is forecasted to increase to 1.9 million people by 2020. The current trend indicates that road traffic accidents will become the 5th leading cause of deaths by year 2030, and thus, there is a critical need to address the above issue (WHO, 2011). It is expected that China, India, Nigeria, Brazil, Indonesia, US, Pakistan, Russia, Thailand and Iran will be the countries at the forefront that contribute significantly to the number of road traffic deaths globally. It is rather alarming that road traffic accidents are recorded as the 5th leading cause of certified deaths in Malaysia and they are the leading 'killer' in China. This global phenomenon has led the United Nations General Assembly to proclaim a Decade Action for Road Safety 2011–2020. Countries all over the world are encouraged to implement

activities according to the five pillars: (1) road safety management, (2) safer roads and mobility, (3) safer vehicles, (4) safer road users and (5) post-crash response.

1.2 Road safety development in Malaysia

The number of road traffic fatalities of Malaysia doubled within a span of 15 years, from 1981 to 1995. The number of fatalities has increased over 150% from 1981 (the first year national fatality statistics are available) to 2012. However, this increase is still considerably lower compared to the increase in vehicle ownership rate, which is 278% in the same period. In contrast, the increase in population is only 106%. The average annual growth of fatalities occurs within these periods: 1981–1983 (13.2%), 1988–1992 (8.1%) and 1993–1996 (10.5%), as shown in Figure 1.1.

In response to the alarming increase in the number of fatalities, a number of agencies under various ministries in Malaysia have been established and integrated to formulate initiatives, strategies, tasks and targets of road safety development. These agencies are directed to undertake road safety measures under their responsibility. The Government of Malaysia has also urged the public and NGOs to participate in road safety campaigns. The Road Safety Cabinet Committee was formed in 1989 to set national road safety targets and since, there is the aim of reduce the number of fatalities to 30%. The Public Works Department (JKR) in the Ministry of Works established the Road Safety Section in 1997. Later, the Ministry of Transport established the Road Safety Department (JKJR) in 2004 and Malaysian Institute of Road Safety Research (MIROS) in 2007.

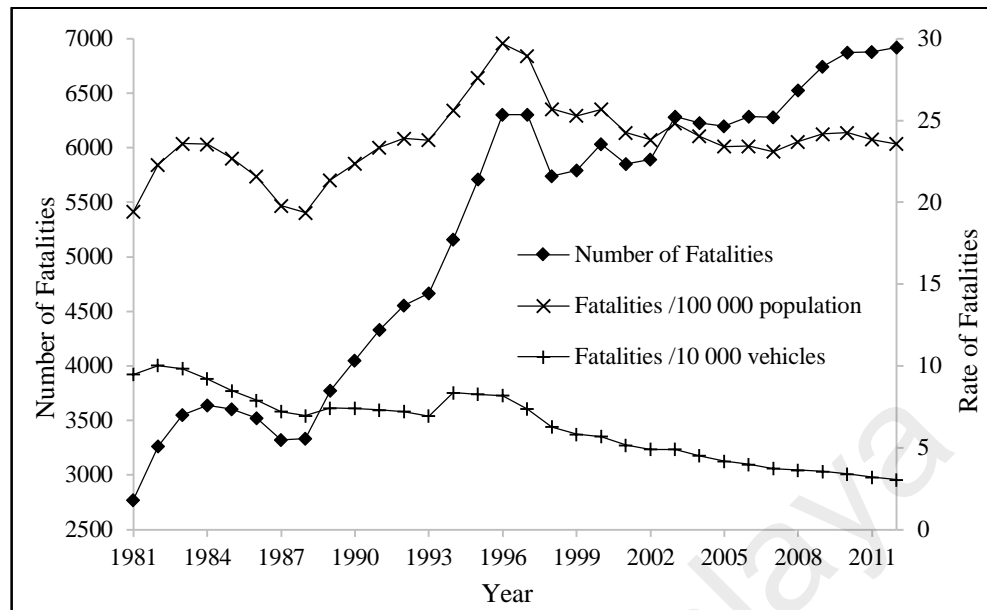


Figure 1.1: Number and rate of road fatalities in Malaysia from 1981 to 2012

Note: The rates of fatalities per 10,000 vehicles are only estimations, rather than the actual values. The data on the number of fatalities were obtained from reports by the Royal Malaysian Police (PDRM, 2013) whereas data on vehicle registration were obtained from reports published by the Road Transport Department of Malaysia (MoT, 2013).

The first national road safety plan was launched in 1996, which includes various initiatives and tasks to improve road safety and achieve the target. These involve the development of a national accident database system, the implementation of five stages of road safety auditing, national blackspot programmes, integrated enforcement and new helmet standards (Radin Umar, 2005). The aim of reducing the number of deaths due to road accidents to 30% by year 2000 was also strengthened as the national target. In his analysis on the first road safety target, Radin Umar (1998, 2005) predicted that the integrated road safety programmes reduce the number of fatalities from 9,127 to 6,389 in year 2000. However, according to police reports, 6,035 deaths were recorded in 2000, which is equivalent to 5.69 deaths per 10,000 vehicles. This indicates that the targeted reduction in the number of fatalities specified in the first national target has been achieved, in accordance with Radin Umar's prediction.

Baguley and Mustafa (1995) presented a detailed calculation of the road safety target for year 2000. According to them, a reduction of 30% means a reduction in the fatality rate from 7.12 deaths per 10,000 registered vehicles (based on the 1989 figures) to 3.14 deaths per 10,000 vehicles by year 2000. This corresponds to a death toll of 2,641 due to road accidents. However, the actual fatality rate per 10,000 vehicles, is considerably higher than Baguley and Mustafa's prediction which is too optimistic. Further examination reveals that the discrepancy between both studies is due to a difference in the interpretation of the national target. Basically, Baguley and Mustafa deduced 30% from 3,773 deaths in 1989 and this results in 2,641 deaths predicted in year 2000. For this reason, Baguley and Mustafa underestimated the number of deaths. In contrast, Radin Umar deduced that the 30% of deaths is from the 'business as usual' scenario forecast.

Hence, the national target, which is to reduce the number of fatalities by 30% in year 2000, is basically achieved. However, the rate of fatalities in Malaysia is still relatively high compared to the rates achieved in developed countries. For this reason, the Road Safety Plan was launched in 2006 in order to reduce the rate of the fatalities to be at par with developed countries. In deference to the first plan, the absolute rates of fatalities were outlined in the Road Safety Plan, rather than the percentage of reduction to be achieved. The road safety targets are as follows: (1) 2.0 deaths per 10,000 registered vehicles, (2) 10 deaths per 100,000 population and (3) 10 deaths per billion vehicle-kilometres travelled by 2010. However, these targets were not achieved since the rate of fatalities in Malaysia was 3.4 per 10,000 registered vehicles and 24.3 per 100,000 population in 2010. These values are still higher than the targets set in the plan. On one hand, the relatively high rates of fatalities indicate that there is a need to enhance road safety strategies. On the other hand, there is a need to define the possible explanations why the national targets are not achieved. It is possible that the targeted rates of fatalities

are too ambitious or there may be faults with the prediction. It is also possible that there is a decrease in the efforts to improve road safety.

In 2012, MIROS published the forecasted number of fatalities in Malaysia up to year 2020. MIROS is a government agency that functions as a one-stop centre to generate and disseminate road safety information. The autoregressive integrated moving average with explanatory variables (ARIMAX) and generalised linear models (i.e. Poisson and negative binomial) were used in the forecast. Based on the report, it is expected that the number of fatalities will continue to increase to 8,760 and 10,716 deaths in 2015 and 2020, respectively. This indicates that annual growth of fatalities in Malaysia is 6.2% within 2011–2020, which is rather significant for a country that aims to achieve the status of a developed country in year 2020. Since it is forecasted that the population in Malaysia will be 32.4 million in 2020 (Dept. of Statistic, 2012), the rate of fatalities will be 33 deaths per 100,000 population in that year, based on MIROS's prediction. With the current road safety developments, it is perceived that the target of 10 deaths per 100,000 population prescribed in the National Road Safety Plan 2006 may not be achieved by 2020 and it may be even worse compared to year 2010. Nevertheless, the forecast by MIROS serves as the basis for the Road Safety Plan of Malaysia 2014–2020. The main objective of the plan is to reduce the predicted 10,716 deaths in 2020 by 50%, i.e. 5,358 deaths (MoT, 2014).

1.3 Problem Statement

The increasing trend in the number of fatalities in a highly motorized country such as Malaysia is rather unusual. Even WHO (2013) estimated a higher number of fatalities in Malaysia in year 2010 (7,085 deaths) than the value officially reported (6,872 deaths). Among the countries with a vehicle ownership rate greater than 0.5, Malaysia has the

highest rate of fatalities per 100,000 population, as shown in Figure 1.2. Moreover, the rate of fatalities in Malaysia is quite astonishing compared to other countries. In general, countries with a rate of fatalities above 20 per 100,000 population will typically have a vehicle ownership rate of 0.4. However, the rate of vehicle ownership in Malaysia is 0.71, which is the typical value for developed countries, and yet the rate of fatalities is relatively high. This leads to the following questions: Are there specific characteristics that cause the rate of fatalities to be different from those in countries with a similar motorization level? Is the methodology used to attain the targeted rate of fatalities suitable for implementation in Malaysia, considering the fact that 47% of the total vehicles are motorcycles? Is validation required for this methodology?

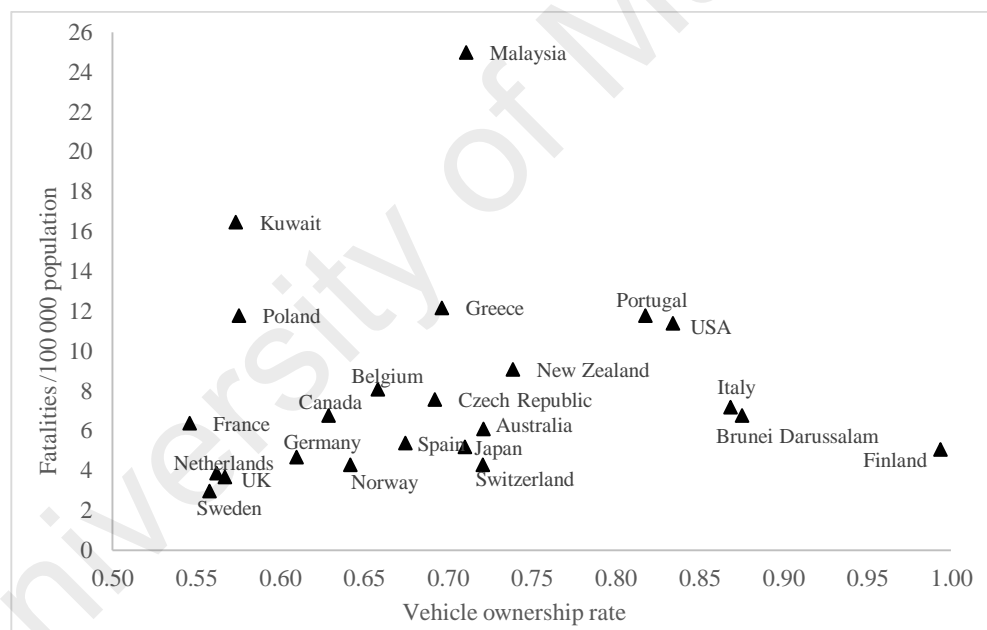


Figure 1.2: Rate of fatalities in countries with a vehicle ownership rate greater than 0.5 in year 2010

Source: Estimated from WHO (2013).

The relatively high proportion of motorcycles however, is not only in Malaysia. The percentage of motorcycles in neighbouring countries in Vietnam, Cambodia, Lao P.D.R, Indonesia and India is even more than 70% of the total traffic. One of the hypotheses that can be related to this situation is that community learning on road safety may not be as

rapid as in the developed countries. At the same time, there are vast improvements in motor vehicle safety and road infrastructures. This in turn, may cause the risk compensation theory to become a reality: improved vehicle safety standards and higher geometric design of roads shall increase the reckless behaviour of drivers.

According to a WHO (2004) report, there are unusual characteristics with regards to the road safety scenario in Malaysia. The report states that Malaysia has experienced a continuous decline in the number of deaths per 10,000 vehicles since 1975, whereas there is a slight increase in the rate of deaths per 100,000 population. Over the same period, there has been rapid growth in motorization and increased mobility among the population in Malaysia. This indicates that the increase in the number of road traffic fatalities is slower in Malaysia compared to the growth of vehicle fleet. However, the number of road traffic fatalities has increased slightly faster in recent years compared to population growth. It shall be noted though that the report does not mention the cause for the current trend in the number of fatalities even though it is stated that more information is required to comprehend how the changes in mobility and safety standards have contributed to such a trend.

In order to demonstrate the characteristics of road safety in Malaysia, a simple analysis is carried out in this study using Smeed's (1949) and Koren and Borsos' (2011) formulae. The analysis is also conducted to determine the relationship between the rate of fatalities per population and rate of vehicle ownership. Smeed's famous model has been used by researchers to predict the number of fatalities since the 1950s – however, the model has been criticised for its lack of accuracy. Other studies have shown that the rates of fatalities predicted using Smeed's model are relatively higher than those obtained from observations (Andreassen, 1991; Jacobs & Cutting, 1986; Jacobs & Hutchinson, 1973; Emenalo et al., 1977; Mekky, 1984; Gharaybeh, 1994; Pramada Valli, 2005).

Koren and Borsos (2011) extended Smeed's model and the new model is more acceptable since it provides a relatively low error when it is fitted to the recent data on global road fatalities. They used the data from 175 countries in year 2010 provided by WHO (2013) in the development of their model and the results are shown in Figure 1.3, which confirm the hypotheses of the opponents to Smeed's model and Koren-Borsos's finding. Firstly, according to the Smeed's model, the rate of fatalities per population will continue to increase with an increase in the rate of vehicle ownership. Secondly, the rate of fatalities will decrease if the rate of vehicle ownership is greater than 0.2.

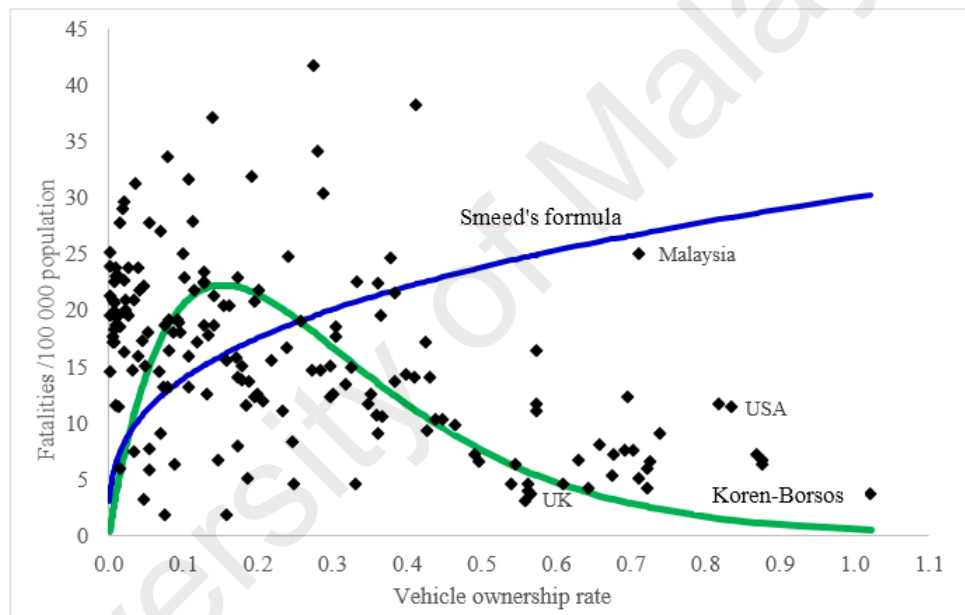


Figure 1.3: Trend of rate of fatalities in 2010 based on Smeed's and Koren-Borsos's models

It is interesting to note that the rate of fatalities in Malaysia in 2010 is relatively close to that predicted by the Smeed's model (Figure 1.3), considering that it is well known that the model was developed based on the data from 20 industrialised countries in 1938. However, the original Smeed's formula has been revised and updated in various studies over the years in order to make it applicable for forecasting the rate of fatalities in the future. In these studies, it is concluded that the formula needs to be modified since there

is a declining trend in the rate of fatalities over the years. This raises the following question: Is it possible that the relationship between the rate of fatalities and motorization level in Malaysia in year 2010 is similar to the scenario in industrialized countries in 1938?

1.4 Objectives and significance of the study

The primary objectives of this study are: (1) to describe the characteristics of road safety in Malaysia, (2) to investigate the impact of road safety measures in reducing the rate of fatalities, (3) to investigate the factors that influencing the rate of fatalities and (4) to develop time series models to predict the rate of road traffic fatalities in Malaysia. The variables that are believed to influence the rate of fatalities in Malaysia are examined. In addition, the major road safety measures taken during the study period are examined to explore their effectiveness in reducing the number of fatalities. In order to reach the objectives, the study conducts the descriptive statistics and develops three time series models, i.e. ARIMA, transfer function-noise and state-space models. Hence, both descriptive and explanatory time series modelling are employed in this study.

The factors and variables that significantly contribute towards road traffic fatalities are identified in this study, which serve as the basis to develop models which will predict the number of fatalities and rate of fatalities up to year 2020. The characteristics of road safety in Malaysia are also identified. Furthermore, the rate of fatalities targeted in the 2006 and 2014 Road Safety Plan are tested to determine whether these targets are achievable by year 2020. The results of the analysis may assist the parties involved in road safety to anticipate the trend of road traffic fatalities in the future and therefore, the necessary measures can be taken if the predicted death toll is beyond expectation. The results of this study can be used as a basis for comparison with the predicted data released by MIROS.

1.5 Research questions

As described above, it can be expected that the number of fatalities due to road traffic accidents in Malaysia will worsen until year 2020 (Rohayu et al., 2012). This phenomenon, however, contradicts with the number of road traffic fatalities in developed countries, in which there is a decreasing trend in the number of road traffic fatalities since the early 1970s, followed by a constant (unvarying) trend in recent years. However, it shall be noted that the fatality statistics extracted from the reports provided by the Malaysian Royal Police have shown only a gradual increase since year 1997. Hence, there is a critical need to rebuild a model to predict the rate of fatalities attributed to road traffic accidents in Malaysia. It is believed that the following questions will be answered with the development of robust forecasting models as well as the descriptive statistics of the road safety data.

- (1) How effective are the road safety measures implemented in Malaysia in reducing the percentage of fatalities? The reduction in the percentage of casualties is compared with the values achieved in other countries.
- (2) Which is the best time series model among the three models developed in this study in order to predict the rate of fatalities in Malaysia?
- (3) Is it acceptable to only use the descriptive/univariate model (without explanatory variables) which is practised extensively in Netherlands?
- (4) What are the significant explanatory variables which affect the number of fatalities?
- (5) Is it possible for the number of fatalities to increase significantly on an ongoing basis considering the fact that most of road safety measures have been implemented?

- (6) Is the major decline of fatalities in 1997–1998 due to the initiatives and measures implemented in Malaysia? Or is it possible that this decline is a consequence of an unexpected regional economic crisis at the time which in turn reduces traffic exposure (vehicle-kilometres travelled or fuel consumption)?
- (7) Is it practical to include only the population and number of vehicles as the explanatory variables (as had been done by MIROS), considering the fact that Smeed's (1949) famous model is refuted in previous studies (e.g. Lassarre, 2001; Yannis et al., 2011a) because it merely considers these two variables in the model?
- (8) Can the number of fatalities targeted in the Road Safety Plan 2014–2020 (death toll: 5,358) be achieved in year 2020?
- (9) Has the risk compensation theory occurred? It is known that improvements in road geometry and the production of vehicles with enhanced safety features lead to an increase in driving intensity (reckless behaviour). Even though enforcements on road safety have been carried out intensively during festive seasons in Malaysia, it is still likely that road safety practices are neglected by road users.

1.6 Methodological framework

In order to address the research issues highlighted in this study, a systematic literature review is first carried out in order to obtain a thorough understanding on road safety modelling. The theories and techniques available in the literature are scrutinized and assessed carefully in order to identify those that are suitable for this study. The descriptive statistics of the data pertaining to road safety that are obtained from national databases that are covered West and East Malaysia as well as of all road categories then determined. Data stratification is carried out since the victims of fatalities are predominantly motorcyclists, with the aim to define the relationship between selected data and the

number of fatalities over the years. The variables which potentially contribute towards road traffic fatalities are then identified.

However, it shall be noted that these variables are only limited to those that can be included in macroscopic modelling. To the best of the author's knowledge, there are no adequate databases that can be used to predict the rate of fatalities on a national scale based on microscopic variables. For instance, the black spot treatment programme through road safety auditing has been proven to be successful in reducing the number of fatal accidents. Difficulties arise, however, when an attempt is made to detail how many black spots remained as since the figures change constantly due to the occurrence of new accidents. For this reason, predicting the total number of fatalities in Malaysia based on the number of remaining black spots as the independent variable is rather impractical at the moment. A more feasible approach is to use the significant explanatory variables proposed by researchers in the field as the groundwork before the inclusion of any other variables. It is also noteworthy that the various models and variables that contribute to fatalities may differ from one country to another. The rate of fatalities is decreasing in developed countries, whereas the trend is increasing in developing countries.

Following this, the explanatory variables that contribute to the number of fatalities need to be defined, and the stepwise regression technique is employed for this purpose. Collinearity is used as the basis to remove any variables from the model. In addition, the time series of various road fatalities since the late 1960s are examined to attain the remarkable turning point(s), followed by a decrease or an increase. The road safety measures implemented years *ex ante* of the major turning point(s) are then examined. The time series intervention model is used to determine the effectiveness of the road safety measures which are believed to bring about the major change in the rate of injuries. In Malaysia, most of the road safety laws have been enacted prior to 1990s. The extensive

and integrated measures have been devised within 1996–1997, following the first National Road Safety Target. After 1997, the policies are focused on the increasing the levels of enforcement, and most of the safety measures implemented are basically ongoing versions of the previous ones. In addition, road safety development relies heavily on road engineering works, media campaigns and education. Road safety is indeed improved by the implementation of the blackspots treatment programme.

Several motor vehicle safety standards have been imposed prior to the availability of national road death toll in 1981, with the exception of the new helmet standard that was mandated in 1996. Hence, it is not possible to trace all of the effects of motor vehicle safety standards at the macro level using time series analysis due to the limited time span of available data. The autoregressive integrated moving average (ARIMA) and transfer function-noise models are used to correlate the variations in the number of fatalities in the year when the safety measures are implemented. Both models have been carried out to assess the effectiveness of the implementation of the safety measures since the 1970s. Moreover, according to Radin Umar (2005, 2007), the following period (1996–1997) is the turning point in road safety development since the safety measures implemented since 1996 results in the most drastic decline in the number of fatalities in Malaysia within 1997–1998. Hence, this intervention is prioritised to be explored. The analysis, however, is taken with precaution since it is possible that the drastic decline is due to the regional economic downturn that leads to a reduction of traffic exposure.

The models described above can be classified into univariate and multivariate time series analysis. In this study, the state-space model is employed for multivariate analysis, whereas ARIMA and intervention analysis through the transfer function-noise models are employed for univariate analysis. A variety of competing models can be developed due to the vast number of variables and statistical software available. Following this, three

statistical selection techniques are employed to determine the best model, namely mean absolute percentage error (MAPE), Akaike information criterion (AIC) or Schwarz's Bayesian criterion (SBC). AIC and SBC are model goodness of fit. The best model is then used to predict the fatalities due to road accidents and provides insight on the road safety scenario in Malaysia. The results also reveal the answers to the questions. These can also be answers to the questions raised in the preceding section. The methodological framework of the methodology used in this study is presented in Figure 1.4.

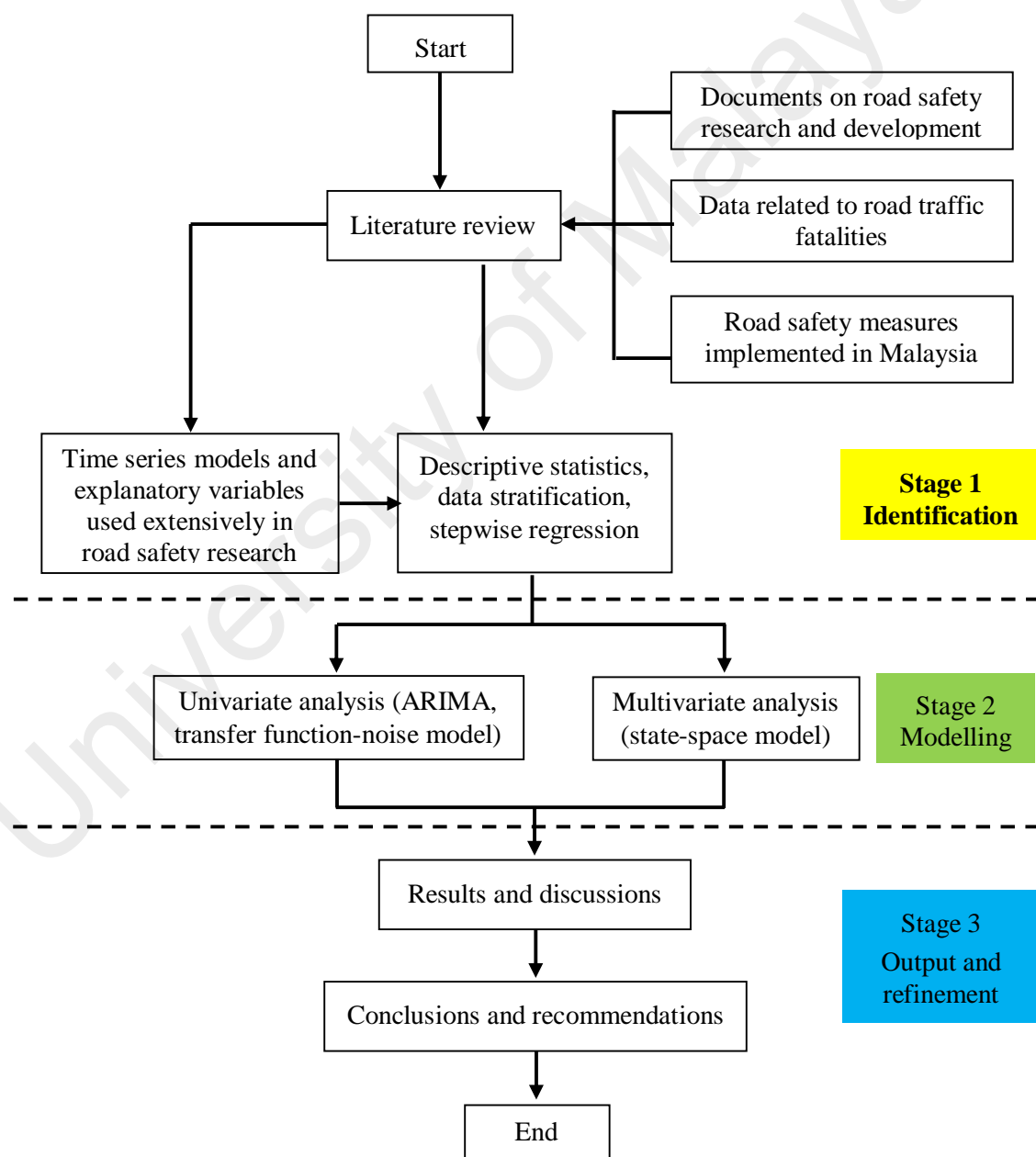


Figure 1.4: Methodological framework of this research

1.7 Organization of the thesis

This thesis is divided into six chapters, including this introductory chapter. The contents of each chapter are summarized in Table 1.1.

Table 1.1: List of thesis chapters and the contents of each chapter

| Chapter | Description | Contents |
|---------|----------------------|---|
| 1 | Introduction | This chapter consists of the background of the study, road safety development in Malaysia, motivation of the study, objectives and significance of the study, research questions and hypotheses, and the conceptual framework of the methodology adopted in the study. The organization of the chapters in the thesis is presented at the end of the chapter. |
| 2 | Literature review | This chapter consists of a review of the history of road safety research and development and a comprehensive review of the literature pertaining to the time series analysis that has been used extensively to predict the rate of fatalities due to road traffic accidents. A review of the statistical techniques used for forecasting is presented at the end of the chapter. |
| 3 | Research methodology | This chapter contains a detailed description of the methodology adopted in this study. The descriptive statistics used for modelling, and identification of the explanatory variables are presented, along with a detailed description of the ARIMA, transfer function and state-space models. |
| 4 | Results and analysis | This chapter consist of the results of this study, beginning with the descriptive statistics which reveal the current road safety scenario in Malaysia. The results of the univariate analysis are presented, which provide insight on the effectiveness on the road safety measures that have been implemented in Malaysia, along with the predicted rate of fatalities up to year 2020 obtained from multivariate analysis. |

Table 1.1, continued

| Chapter | Description | Contents |
|---------|--|---|
| 5 | Discussion and implications of the study | The results of this study are discussed in detail in this chapter and comparisons are made with the findings of other researchers. The research questions and hypotheses presented in the first chapter are addressed and the implications of this study are discussed. |
| 6 | Conclusions and recommendations | The key findings of this study are summarized in this chapter and recommendations are given to improve the road safety scenario in Malaysia in the future. The recommendations for future work are also presented in this chapter. |

1.8 Chapter summary

In Malaysia, with the current road safety developments, the target of 10 deaths per 100,000 population prescribed in the National Road Safety Plan 2006 may not be achieved by 2020 and it may be even worse compared to year 2010. It is also perceived that the target of the latest Road Safety Plan of Malaysia 2014–2020 that is the main objective is to reduce the deaths in 2020 by 5,358 deaths may not be achieved.

Among the countries with a vehicle ownership rate greater than 0.5, Malaysia has the highest rate of fatalities per 100,000 population. In general, countries with a rate of fatalities above 20 per 100,000 population will typically have a vehicle ownership rate of 0.4. However, the rate of vehicle ownership in Malaysia is 0.71, which is the typical value for developed countries, and yet the rate of fatalities is relatively high. It is interesting to note that the rate of fatalities in Malaysia in 2010 is relatively close to that predicted by the Smeeds model, considering that it is well known that the model was developed based on the data from 20 industrialized countries in 1938.

The primary objectives of this study are: (1) to describe the characteristics of road safety in Malaysia, (2) to investigate the impact of road safety measures in reducing the rate of fatalities, (3) to investigate the factors that influencing the rate of fatalities and (4) to develop time series models to predict the rate of road traffic fatalities in Malaysia. The factors and variables that significantly contribute towards road traffic fatalities are identified in this study, which serve as the basis to develop models which will predict the number of fatalities and rate of fatalities up to year 2020. The results of the analysis may assist the parties involved in road safety to anticipate the trend of road traffic fatalities in the future and therefore, the necessary measures can be taken if the predicted death toll is beyond expectation.

There are nine research questions arise in this study. The questions will be answered with the development of robust forecasting models as well as the descriptive statistics of the road safety data. The variables which potentially contribute towards road traffic fatalities are identified. However, it shall be noted that these variables are only limited to those that can be included in macroscopic modelling. There are three time series models employed in this study, i.e. the autoregressive integrated moving average (ARIMA), transfer function-noise and state-space models. ARIMA and transfer function-noise models are used to correlate the variations in the number of fatalities in the year when the safety measures are implemented. The state-space model is employed for multivariate analysis, whereas ARIMA and intervention analysis through the transfer function-noise models are employed for univariate analysis. The best model is then used to predict the rate of fatalities due to road traffic accidents and provides insight on the road safety scenario in Malaysia.

CHAPTER 2: LITERATURE REVIEW

A detailed review of the literature related to road safety research and time series modelling is presented in this chapter, beginning with a brief history of road safety research and development. The methodologies, key findings and recommendations of other studies are referred to in order to develop the conceptual framework and formulate the research questions in this study. Road safety research in Malaysia is included in order to gain insight on the characteristics of the road safety scenario in this country. The methodologies used to forecast time series data are presented in the latter part of this chapter.

2.1 Road safety research development in early years

The escalating number and rate of fatalities over the years has increased awareness concerning road safety among researchers, practitioners and policy-makers ever since two centuries ago. The authorities in England imposed penalties for offences related to carts and other vehicle operations on the roads in 1835 (Anonymous, 1835). Penalties were imposed for intrusions on motorways and cycling on footpaths. In the early 18th century, the increasing death tolls due to road traffic accidents has raised awareness regarding road safety. Byrne (1920), a medical doctor who worked as an assistant registrar of records at the City of New York Department of Health, urged that radical measures of restriction should be taken in order to protect lives on the roads, and called for a campaign to devise standard regulations with real penalties. He proposed that pedestrians, particularly children, should be given the necessary education concerning road safety. Butterworth (1932) suggested that the problems related to road accidents are not only related to the measures that can be taken to prevent accidents, but also how to bear the consequences

in the event of accidents. These concerns were expressed through the mass media available at the time, raising awareness among the general public.

In the early years, the main goal of highway authorities is to develop internal scientific research and countermeasures in order to reduce road traffic accidents. These measures were then implemented in accident-prone areas. However, the road safety measures implemented at the time obstructs traffic movement and are therefore, criticized by the public. Newspapers and magazines became the platform for the public to express their dissatisfaction and make requests regarding road safety developments. This debate has drawn the attention of road safety experts and some published articles and books regarding the subject. One of the key works at the time is the work of Halsey (1941). The road safety measures were improved based on the feedback received from the public, and this is the way road safety research and development is carried out during the early years. The history of road safety research and development begins in 1940s, which marks the turning point in road safety research.

The World Health Organization (WHO) has given much focus on road traffic fatalities since the early years of road safety research, as highlighted in a report by Norman (1962). This report highlights that motor vehicle collisions constitute the largest death tolls and tend to be the most severe among the various types of accidents. The causal factors are also specified in the report, along with the rate of fatalities from motor vehicle accidents per 100,000 population in selected countries. The rate of fatalities are presented according to age group and gender. Haddon (1968) may be the first to introduce a systematic approach to identify the causes of road accidents, which is known as the Haddon Matrix. This matrix shows that the main components of road traffic accidents that occur at motorways are humans, vehicles and the environment of the motorways. These components are correlated with the motorway accidents during three phases: pre-crash,

in-crash and post-crash. The use of Haddon matrix in reducing the rate of casualties was also accepted by the WHO, which is presented in a report on the prevention of road traffic injuries. This report was edited by Peden et al. in 2004. However, Haddon (1968) did not mention the significance of each component on motorway accidents.

The decrease in the death tolls due to road traffic accidents in developed countries in the early 1970s is one of the crowning achievements in road safety research. Other key achievements in road safety research and development include the formulation of the risk compensation theory, the identification of new causal factors and the development of various techniques to analyse and forecast road traffic fatalities.

2.2 Seminal works in road safety research and development

It shall be noted systematic road safety research has been conducted in the UK since 1947, which forms the basis for the development of regulations, standards and legislations (Smeed, 1952). One of the influential works in road safety research is by Smeed (1949), who used data on road traffic fatalities from 20 countries in 1938 in order to develop a model. He discovered that road traffic fatalities are the product of the number of registered vehicles and the number of population. This model was initially accepted and implemented in a large number of studies (Glanville, 1954; Jacobs & Hutchinson 1973, Jacobs 1982, Mekky 1985). However, the key findings of this study were criticized by some researchers in the late 1980s (Andreassen 1985, Adams 1987, Broughton 1988). Even though Smeed used relevant datasets from various countries, the variable chosen to measure the exposure to risk in Smeed's model is not satisfactory for several reasons (Bergel-Hayat, 2012).

A detailed explanation on Smeed's controversial findings is given in the COST 329 report which is released by EC (2004). It is stated in this report that according to Smeed's

law, a decrease in the number of fatalities is not possible with an increase in the number of registered vehicles and number of population. With this law, one can expect a non-linear increase, but certainly not a decrease, which is contradictory to the scenario in recent years. Jacobs and Hutchinson (1973) further analysed Smeed's model and found that with this model, the number of fatalities per motor vehicle tends to decrease with an increase in vehicle ownership – however, the actual decrease is faster than expected. In recent years, however, some researchers have used the modified Smeed's model. Pramada Valli (2005) discovered that Smeed's formula and Andreassen's equations perform well in estimating the number of road accidents in seven metropolitan cities in India. Koren and Borsos (2011) modified Smeed's formula to produce a model that fit wells with the relationship between vehicle ownership and rate of fatalities based on recent data.

With the exception of Smeed's (1949) model, there are no remarkable theories regarding modelling road traffic fatalities at the macro level in the first half of the 20th century. Glanville (1954) applied Smeed's formula using more recent data. This work, however, may be the first to highlight that the number of road traffic fatalities can decrease with an increase in the number of vehicles. The findings of this study also reveal that road safety is associated with human, vehicle and road factors, which made researchers inclined towards modelling road safety at the micro level. Hence, researchers began to focus their attention on the causal factors of road traffic accidents and associated these factors with the probability of accidents rather than predicting the number of casualties in a nation. This is possibly due to the complexity of road safety modelling and even Smeed's formula has been shown to produce inaccurate predictions. The lack of data at the national level may also be the reason. Consequently, researchers tend to follow Smeed's method on macroscopic modelling.

According to Hakim et al. (1991), studies pertaining to microscopic modelling are of a limited scope – dealing only with a road section, a short time period, or a sample of drivers. For the purpose of research or policy-making, analysis of a single intervention variable may be conducted by means of an impact study, including pre-observations and post-observations.

The growing debate on Smeed's model has led to the formulation of various theories and hypotheses concerning road safety. One of interesting theories in this area was forwarded by Minter (1987), which was then further improved by Koornstra (1987). They concluded that the decrease in road fatalities is attributed to the drivers or community learning. Hence, researchers felt compelled to determine the variables associated with the learning process of drivers, which are not easy to pinpoint.

Sabey (1991) attempted to determine the causal factors of road accidents and found that some estimates were based on analytical measures, which was not practical for use in areas such as education and publicity. For this reason, a large number of studies employed univariate time series models or extrapolation techniques which exclude all causal variables – rather, these models only show the changes in the variables with respect to time. Moreover, it has been observed that in most large jurisdictions, there is a period in which the number of fatalities increases, reaches a peak point, followed by a period of decline and stabilization. However, all causal variables such as population and traffic exposure are found to increase. Hence, this is another reason for shunning causal models in predicting road traffic fatalities in large jurisdictions (Hauer, 2010). In addition, this approach may not be suitable for forecasting, particularly in less developed countries which have yet to experience a single peak in the absolute number of fatalities. Despite the criticisms, there are still researchers who support Smeed's model in recent years. Koren and Borsos (2011) found that Smeed's formula describes the increase in fatalities

reasonably well, up to a level of 0.2–0.3 vehicles/person ownership. However, the prediction produced by Smeed's formula is rather pessimistic above this level. Fortunately, the number of road traffic fatalities tends to decrease in reality. Hence, they incorporated data from 139 countries published in year 2007 and a new approach to improve Smeed's formula. They successfully produced a model that shows a good fit for the relationship between the rate of fatalities and vehicle ownership.

2.3 Explanatory variables appear to affect fatalities

In the early stages of road safety research at the macro level, the variables investigated involve alcohol consumption, human variables, seat belt usage, speed limits, traffic signals, type of vehicles, vehicle inspection, annual average daily traffic (AADT), type of road users, vehicle age, driver age and gender, as well as certain demographic and transport factors (Bjerver & Goldberg, 1951; British Medical Association, 1960; Cassie & Allan, 1961; Cohen et al., 1958; Glanville, 1951; McFarland et al., 1955; Tourin, 1958; Smeed, 1961; Webster & Newby, 1964; Clyde, 1964; Malo, 1967; Norman, 1962; Garrett & Braunstein, 1962; Huelke & Gikas, 1968; Colton & Buxbaum, 1968; Jorgensen, 1969; Mackay, 1969; Joksche & Wuerdemann, 1970; Carroll, 1973; Trichopoulos et al., 1975). Some researchers have also attempted to determine the economic cost of traffic accidents – one of these is the work of Reynolds (1956).

Even though road safety research has been carried out extensively since the 1970s, only a few studies are focused on macroscopic models. Researchers have used a variety of variables that are expected to influence the number and rate of casualties along with various modelling techniques in order to gain understanding on the road safety scenario. Researchers have also identified a number of variables that are correlated with the number and rate of road casualties in the 1980s. These variables can be categorized into several

factors: (1) humans factor: the proportion of the population, BAC, age and gender, (2) road environment factor: type of road, the condition of the road, length of road networks, time of accident, speed limits and traffic control at junctions, and (3) vehicle factor: the number of registered vehicles, seat belt usage, type of vehicle, vehicle inspection, vehicle ownership and vehicle age.

A comprehensive list of the potential explanatory variables related to road traffic fatalities can be found in the following studies. Hakim et al. (1991) presented a critical review of state-of-the-art macro models for road accidents. This review is intended to identify and establish the significance of policy and socioeconomic variables which affect the level of road accidents. The variables that appear to affect the number of fatalities or injuries are: (1) vehicle miles travelled (VMT), (2) vehicle population, (3) income (in various forms), (4) percentage of young drivers, (5) intervention policies such as speed limits, (6) periodic vehicle inspection, and (7) minimum alcohol drinking age. Page (2001) used the following variables to compare the road mortality of OECD countries: (1) population, (2) motorized vehicle fleet per capita, (3) percentage of urban population, (4) percentage of population aged 15–24 years, (5) percentage of population who are active and employed, (6) consumption of pure alcohol per capita (litres), and (7) percentage of buses and coaches in the motorized vehicle fleet.

Sauerzapf et al. (2010) examined the explanatory variables that have been assessed in previous studies which include: (1) density of the road network, (2) number and type of vehicle per kilometre of road, (3) number of vehicles per 1,000 population, (4) exposure to risk (kilometres travelled per year), (5) quality of available medical services, (6) behavioural factors (seat belt use, speeding and alcohol consumption), as well as national economic factors such as: (7) the percentage of population who are unemployed, (8) Human Development Index and (9) Gross Domestic Product (GDP) per capita.

The independent variables included in the WHO's (2013) model are: (1) ln (GDP), (2) ln (vehicles per capita), (3) road density, (4) national speed limits on rural roads, (5) national speed limits on urban roads, (6) health system access, (7) alcohol – apparent consumption, (8) the number of population working, (9) percentage of motorcycles, (10) corruption index, (11) national policies for walking/cycling, and (12) population.

2.3.1 Socioeconomic and demographic indicators

Yannis et al. (2011b) examined the trends of road traffic fatalities in several EU countries using the temporal evolution of elementary socioeconomic indicators, namely motorized vehicle fleet and population at the national level. Even though road traffic fatalities correspond to a significant part of the traffic, mopeds are not included in the motorization rate in Greece so that the data are comparable with those from other countries. Furthermore, different countries achieve motorization rates at different (and sometimes distant) temporal landmarks. Some of these countries exhibit a breakpoint within a narrow range of motorization rate values, which implies similar social and economic conditions and/or similar road safety cultures. This range is different for certain sub-groups among the countries examined, which suggests that some of the groups may be of geographic and socioeconomic context. The breakpoint of road traffic fatalities of these countries occurs at a motorization rate of 125 to 325 vehicles per 1,000 inhabitants.

2.3.1.1 GDP and household income

In the mid-1980s, a number of researchers have developed macroscopic models which include variations of the models that existed at the time. These models even address the changes in economic factors that are believed to affect road traffic accidents and fatalities (Partyka, 1984; Wagenaar, 1984; Hedlund et al., 1984; Wintemute, 1985). Van Beeck et

al. (1991) discovered that higher income is associated with lower mortalities. Scuffham and Langley (2002) examined the traffic crashes of New Zealand, and they concluded that an increase in income is associated with a short-term decrease in risk but increases the exposure to accidents (distance travelled) in the long term. Kopits and Cropper (2005) used income per capita in their model to project traffic fatalities and stock of motor vehicles until 2020. Richardson and Shaw (2009) analysed the relationship between the income of different states (GDP) and rate of fatalities.

Kopits and Cropper (2005) examined the relationship between traffic fatality risk and income per capita, and used this relationship to forecast road traffic fatalities according to geographical regions. The natural logarithm of fatalities per population (F/P), vehicles per population (V/P) and fatalities per vehicles (F/V) were expressed as spline (piecewise linear) functions of the logarithm of the actual GDP per capita (measured in 1985 international prices). The region-specific time trends from 1963 to 1999 were analysed using linear and log-linear models. These models were used to project traffic fatalities and the stock of motor vehicles in 2020. Based on the results, it is expected that the global road deaths will increase to over 1.2 million in 2020. The study appears to be the first in which the researchers represent the relationship between the risk of road traffic fatalities and income per capita using the Kuznets curve. They applied the Kuznets curve to a large international dataset and projected the future road traffic fatalities using this curve (McManus, 2007). However, Gaudry and Lapparent (2013a) criticized the model developed by Kopits and Cropper, since they did not address the 1972–1973 turnaround – let alone seeing it as a turning point.

Bergel-Hayat (2012) conducted a systematic review and highlighted that risk exposure can be assessed based on the economic activity (for example, measured with respect to the household income). The WHO (2013) included GDP and the number of people

working as the independent variables of the regression models. Yannis et al. (2014) investigated the association between changes in GDP with the changes in the annual road traffic mortality rates in 27 European countries. They concluded that the decrease in road traffic fatalities cannot be solely justified by policy efforts, and the decrease can be partially attributed to the global economic recession which affects both the economy and mobility in most of the European countries. In general, an annual increase of GDP per capita leads to an annual increase of mortality rates, whereas an annual decrease of GDP per capita leads to an annual decrease of mortality rates.

McManus (2007) also examined some of the recent theories which explain the increase in the rate of road traffic fatalities in developing countries. One of the main explanations is that the number of road traffic fatalities per capita follows the Kuznets curve (an inverted U-shaped pattern) with an increase in the income per capita. Based on this trend, it can be expected that the increase in road traffic fatalities per capita in developing countries will reach a certain peak, followed by a decline thereafter. According to Zhang et al. (2011), the road fatality rate per population follows an inverted-U-shaped trend (known as the Kuznets curve) with the level of economic development (GDP per capita).

Traynor (2008) found that there is a significant interaction between the income per capita and the percentage of highway vehicle miles travelled (VMT) in Ohio, indicating that there is a non-linear correlation between the income per capita and rate of fatalities. Based on this correlation, it can be deduced that there is an inverse relationship between the rate of fatalities and VMT when the VMT is low, whereas there is direct relationship between the rate of fatalities and VMT when the VMT is high.

2.3.1.2 Vehicle ownership

Page (2001) concluded that a 10% increase in the motorized vehicle fleet leads to a 1.6% decrease in the number of fatalities. The results showed that the increasing trend in car ownership in many countries will decrease the number of road traffic fatalities. Law et al. (2005) projected the rate of vehicle ownership in Malaysia for year 2010 in order to predict the rate of road accident deaths. Based on the results, they expected that the rate of road accidents will decrease to 4.22 per 10,000 vehicles in year 2010. Richardson and Shaw (2009) found that the increase in the number of vehicles in a Member State leads to an increase in traffic fatalities. However, states with higher incomes (GDP) have lower rates of fatalities. Yannis et al. (2011b) examined the trend of road traffic fatalities in several EU countries from the temporal evolution of elementary socioeconomic indicators at the national level. The elementary socioeconomic indicators are motorized vehicle fleet and population.

Haight (1984) attempted to answer the question, ‘Why does the rate of fatalities per capita decrease?’ He highlighted that there is an indication of travel saturation beginning from about 1979. Shefer (1994) claimed that an increase in travel density beyond a certain point will result in a decrease in road fatalities, *ceteris paribus* (assuming that all other factors remain constant), once road congestion sets in.

2.3.1.3 Unemployment rate

Farmer (1997) used the employment rate and the number of new car sales to analyse fatality trends. According to the predictions produced by his model, the number of motor vehicle fatalities will increase by 12 for every 100,000 decrease in the number of unemployed people for a particular month, given that there are no variations in the vehicle miles travelled. The DRAG framework of Gaudry (2002) also includes an economic

variable (unemployment) in forecasting the casualties. Gerdtham and Ruhm (2006) used aggregate data for 23 OECD countries within 1960–1997 in order to examine the relationship between macroeconomic conditions and fatalities. The results showed that a 1% decrease in the national unemployment rate is associated with a 2.1% increase in deaths due to motor vehicle accidents. The findings of Krystek and Żukowska (2007) confirm the hypothesis that an increase in the unemployment rate will result in a decrease in the number of fatalities.

2.3.1.4 Economic growth

In general, researchers believe that the number of fatalities will decrease with an increase in economic performance (Partyka, 1984; Wagenaar, 1984; Hedlund et al., 1984; Wintemute, 1985). Rizzi (2011) believed that the economic recession of 1998 had an important effect since the number of road accidents declined after 1982 when Chile faced severe recession and disagreed the notion that the action of CONASET on its own decreases road traffic accidents, as asserted by the Commission for Global Road Safety in the late 1990s. This is due to the fact that there are other factors which also contribute to the reduction in road traffic accidents since 1999, but the contributions of these factors are not easily determined. For example, the improvements in interurban roads and the production of cars with higher quality during the Asian Crisis would have definitely played a key role in reducing the number of road traffic accidents. More importantly, it is believed that the economic recession of 1998 has an important effect since the number of road traffic accidents in Chile declined after 1982 – the year when the country is afflicted with severe recession. The use of speed cameras was suspended while the maximum speed limits were increased in 2002, and the economy began to improve in 2003. It was observed that the number of road traffic accidents began to rise in 2003 and the relatively

low number of accidents from 2001 to 2002 is never seen again. It appears that the number of road traffic fatalities has stabilized in Chile.

2.3.1.5 Fuel price

Fridstrøm et al. (1995) examined the monthly number of accidents along with candidate explanatory factors at various counties in Denmark, Finland, Norway, and Sweden. They discovered that the number of road accidents is influenced by random variations as well as a variety of systematic, causal factors. The relationship between exposure (approximated by the gasoline process) and injury accidents appears to be nearly proportional – however, the relationship is less proportional for fatal accidents or death victims. Both the randomness and exposure account for 80 to 90% of the observable variations in the datasets. In Sweden, it is predicted that the number of fatal accidents increases by 0.64% for a 1% increase in gasoline sales. It is interesting to note that the accident toll can be reduced significantly without a decrease in traffic volume, considering that it is one of the crucial systematic determinants.

Chi et al. (2013) used time geography theory to explain that the effect of gasoline prices on traffic crashes occurs through reduced travel and changes in the driving behaviour of drivers in Mississippi. They observed that higher gasoline prices lead to higher road traffic safety. However, it is interesting to note that higher gasoline prices discourage driving and changes driving behaviour such that the drivers reduce fuel consumption and vehicle speed, and avoid driving under the influence of alcohol. All of these contribute to the increase in road safety.

2.3.2 Age, gender categories and urbanization

According to the US DOT report (cited by Cobb and Coughlin (1998)), *there is a well-established U-shaped relationship between driver age and per-mile accident involvement rates, with older drivers having greater rates of accidents, injuries, and fatalities than middle-aged drivers*. Bossche et al. (2007) observed a similar trend in their study. They conducted a road safety analysis for different age and gender categories of road users, and discovered that there is U-shaped relationship for both males and females, which indicates that the risk of fatalities is higher for younger and elderly people. They also observed that the risk is lower for females compared to males, even though the difference is rather small for the first age category.

Fergusson (2003) concluded that risky driving behaviours are common among young people, particularly young males who are prone to externalizing behaviours such as substance abuse, crime and affiliations with deviant peers. Oster and Strong (2013) stated that in 2009, drivers from these age groups (16–20 years and 21–24 years) make up the highest rate of fatalities compared to drivers from other age groups. However, the rate of fatalities of these age groups decreases between 1980 and 2009 in the US by 2.3 and 2.4% respectively. In contrast, older drivers who initially had a lower rate of fatalities have a higher rate of fatalities in 2009. Similarly, women, whose highway fatality rate is initially well below that of men from all age groups, have a higher highway fatality rate in 2009. This is likely due to their increased share of vehicle miles travelled which is attributed to the increase of women participation in the labour force over the years.

Page (2001) concluded that a 10% increase in the urban population leads to a 4.1% decrease in the number of fatalities. The results showed that the increasing trend in urbanization in many countries will decrease the number of road traffic fatalities.

2.3.3 Weather and time variations

Researchers have associated variations in weather conditions with road traffic accidents (Scott, 1986; Fridstrøm et al., 1995; Pisano et al., 2008). The impact of weather conditions was examined by Scott (1986), who concluded that an increase in the amounts of rainfall is associated with an increase in the number of road traffic accidents. Fridstrøm et al. (1995) observed that weather conditions have a significant impact on the number of accidents even though the direction of impact may seem counterintuitive in some cases. Bossche et al. (2004) found that weather conditions have a significant impact on traffic safety. Hermans et al. (2005) analysed the impact of meteorological conditions on road safety, and the results showed that the number of days with precipitation is a variable which significantly affects the number of fatalities on motorways. In another study, Hermans et al. (2006) found that precipitation and thunderstorm are the factors that affect the number of light-injury accidents.

Pisano et al. (2008) agreed that weather conditions have a significant effect on road safety since one quarter of all road traffic accidents that occur on public roads in the US are weather-related. Yannis and Karlaftis (2009) attempted to determine the relationship between the decrease in the amount of rainfall with the total number of accidents and fatalities. Their findings, however, are contradictory to the findings of others. They observed that an increase in the amount of rainfall decreases the total number of accidents and fatalities as well as the number of pedestrian accidents and fatalities. This may be attributed to the safety offset hypothesis which results from a change in driving behaviour, whereby drivers generally become more cautious and they will reduce their vehicle speeds while driving in the rain.

Sukhai et al. (2011) examined 20 explanatory variables related to traffic, weather and the number of public holidays. The results showed that alcohol and fuel sales, coupled

with school holidays and the number of fatalities from two preceding weeks are significant predictors of weekly road traffic fatalities in South Africa, accounting for roughly one-third of the weekly variations in road traffic fatalities after adjustments have been made for at-risk population sizes.

Bjørnskau (2011) categorized the data on road traffic risks according to the day of the week and the time of day, and the results showed that the risk of injuries is extremely high for car drivers and passengers on Saturday nights/Sunday mornings. Bossche (2005) discovered that the trend and trading day are the significant variables which influence the number of people killed or the number of people seriously injured due to road traffic accidents in Belgium.

2.3.4 Medical care facilities

Jacobs and Cutting (1986) determined that there is a significant relationship between fatality index and the number of doctors per capita. Van Beeck et al. (1991) found that traffic density and the availability of advanced trauma care (neurosurgery and computerized tomography) in the region are the most important predictors of the differences in regional mortalities. Both of these variables show an inverse relationship with the number of fatality cases. It is highly likely that higher traffic density leads to a shift towards less severe injuries. The availability of advanced trauma care is important for early diagnosis and treatment of head injuries. Noland (2003) concluded that medical care reduces traffic-related fatalities in developed countries.

According to Bener (2009), the rate of fatalities is associated with health, social and economic parameters which include: (1) population per physician, (2) population per hospital bed, (3) school-age population attending schools, (4) vehicle ownership and (5) gross national product per capita.

2.3.5 Geographical factors and road geometry

Jones et al. (2008) implemented a geographical approach to analyse the number of fatalities, the number of serious casualties and the number of slight casualties due to road traffic accidents in local districts in England and Wales. They examined a variety of potential explanatory variables such as the number and characteristics of the population, traffic exposure, road length, curvature and junction density, land use, elevation and hilliness, as well as climate. The results showed that the use of a geographical approach in road traffic crash analysis helps in identifying contextual associations which conventional studies on individual road sections would have missed.

Noland (2001) examined the effect of various highway improvements on road traffic fatalities and injuries. The improvements include the total lane miles of capacity, the total average number of lanes according to functional road category (interstates, arterials and collectors), lane widths and the relative balance of various road categories within each state in the US. The results strongly refute the hypothesis that improvements in engineering design have been beneficial in reducing the total number of fatalities and injuries. In addition, the results showed that demographic changes in age cohorts, increased seat belt use, and increases in medical technology contribute significantly to the overall reduction in fatalities.

Wang et al. (2013) also reviewed recent studies on road safety and the effects of various factors on road safety, particularly those which are related to traffic characteristics (traffic speed, density, flow and congestion) as well as road characteristics (road geometry and infrastructure). The review is focused on car accidents that occur on major roads such as motorways. They concluded that some factors require further investigation (e.g. speed, congestion and road curvature) since the findings of the studies are inconclusive. Muhammad Marizwan et al. (2013) suggested that there is a need to establish a proper

and systematic road geometry and traffic census inventory in order to develop better accident prediction models for Malaysia in the future.

Oster and Strong (2013) stated that the challenges in evaluating highway safety policies and performance involve a multitude of factors which are not limited to driver behaviour, vehicle design and traffic engineering. Changing demographics and travel patterns have both contributed to the reduction in the overall highway fatality rates. Public policies also play an important role. However, it is rather difficult to evaluate how effective an individual policy has been in reducing the overall highway fatality rates – let alone to identify which policies are really effective in improving road safety. Ongoing urbanization in the US is a major factor. The urban highway fatality rate is only 36% of the rural fatality rate in 2009. Since 1980, the urban highway fatality rate has decreased by 72% whereas the rural fatality rate has decreased by 56%.

2.3.6 Proportion of motorcycle and public transit user

Garbacz (1989) investigated the relationship between the proportion of motorcycles and the number of road traffic fatalities in Taiwan, in which the proportion of motorcycles is very high. They concluded that motorcycles are positively correlated to the number of fatalities, even though to a smaller degree than expected. According to Mohan (2002), vulnerable road users make up the highest portion of traffic in low-income countries. For this reason, vehicles, roads and road environments should be designed to cater the safety needs of these road users. Even though a complete solution is unavailable, new policies and designs can be introduced in order to mitigate this problem. Besides the crashworthiness of vehicles, it is believed that transportation planning, exposure control, intelligent separation of non-motorized traffic on major roads, and traffic calming play a more important role in reducing road traffic fatalities.

In a recent study, Bjørnskau (2011) updated the road traffic risk figures in Norway. The estimates for 2009 and 2010 showed that there is reduction in road traffic risks compared to previous years. The risk levels have decreased over time for all road user groups in Norway – however, the decrease is most pronounced for motorcyclists. The WHO (2013) has also included the percentage of motorcycles as one of the independent variables in its regression models.

Yannis et al. (2011b) made the following conclusions based on a detailed review of road safety research: (1) urban sprawl can be a risk factor in road traffic fatalities, (2) oil prices can affect vehicle travel which in turn affects traffic safety, and (3) road traffic fatalities per capita decrease with an increase in public transit and non-motorized travel activity or mild transport modes such as walking and cycling.

2.3.7 Other explanatory variables

In terms of the exposure to risks, Brüde (1995) developed a model that uses traffic mileage index as the main variable to predict the number of fatalities in Sweden until year 2000. It has been claimed in previous studies that the attitude of drivers is the most important factor which influences the number and rate of road traffic fatalities – however, it is also the most difficult variable to measure since attitude is a subjective quantity. Bester (2001) used the Human Development Index (HDI) as one of the explanatory variables. The final model includes passenger car ownership, HDI, and the percentage of other vehicles as the variables.

According to Raftery et al. (2011), both Australian and international sources indicate that single vehicle crashes, particularly those due to loss of vehicle control, account for the majority of heavy vehicle crashes. The statistics published by the National Transport Insurance (NTI) also indicate that 32 and 10% of heavy vehicle crashes are attributed to

vehicle speed and fatigue, respectively. However, the statistics showed that 42% of heavy vehicle crashes occur due to the combination of fatigue and vehicle speed. Whilst, in recent trend, Mohammadi (2009) conducted a study in Iran and found that the use of mobile phones while driving increases the risk of road crashes.

2.4 The effect of fuel shortages on road safety in 1973

In the early 1970s, an energy crisis occurred in industrialized countries due to the Middle East oil embargo. The fuel shortages have led the authorities, specifically in the US, to take some drastic measures (known as energy conservation) in order to reduce fuel consumption including fuels consumed by motor vehicles. The Congress passed emergency energy conservation legislation which mandates a national speed limit of 55 miles per hour (mph), which is lower than the speed limits at the time. Each state in the US sets their own speed limits and the typical speed limit for most motorways is 65 and 70 mph (Hoskin, 1986). Researchers believed that fuel conservation has a significant impact on road safety. Haight (1994), however, noted this achievement as ‘serendipitous piggybacking’.

The shortage of gasoline, reduction in the maximum speed limits and changing driving patterns in most states results in a decline in the overall road traffic fatalities. Some states experienced a reduction in road traffic fatalities up to 80% due to a reduction in speed limits. (Tihansky, 1974). Chu and Nunn (1976) studied the impact of fuel shortage in detail and the results indicate that the reduction in road traffic fatalities in 1974 is mainly due to the energy crisis. The factors which largely contribute to this reduction are reduced travel, permanent daylight time savings, and a decrease in driving speeds. Each of these accounts for 29, 8 and 3% in the reduction of road traffic fatalities, respectively. The remaining 24% is attributed to other factors such as a reduction in the average occupancy,

changes in day-night travel, changes in the types of roads, and an increase in safety belt use. A number of studies focused on the impact of fuel shortage have shown the relationship between reduced speed limits and the number of road traffic fatalities.

Baruya (1997) highlighted that the effect of speed on the number of accidents in various countries is most apparent during the oil crisis in 1973. Hakim et al. (1991) compiled some US reports and summarized that the overall fatality rate driven per mile on all US roads decreased by 15% after the introduction of reduced speed limits in early 1974. This is 'an unprecedented decline in a single year'. The rate of fatalities decreases by 32% on interstate highways. According to NHTSA (1998b), the number of traffic fatalities decreases throughout the nation from 54,000 in 1973 to 45,000 in 1974 following the enactment of energy conservation measures. This decrease was attributed to a decrease in speed limits.

2.5 The decrease in death tolls due to road traffic accidents in 1972–1973

Industrialized countries experienced a sudden decrease in the absolute number of road traffic fatalities in 1972–1973. Among the countries that experienced a pronounced decrease in road traffic fatalities are Belgium, Great Britain, Denmark, Finland, France, Germany, Israel, Italy, Japan, the Netherlands, Norway, New Zealand, Switzerland, Sweden and the US (Adams, 1982). A number of ideas have been proposed by researchers in relation to this drastic decline. In the US, it is believed that energy conservation strategies (such as imposing lower speed limits) is the main factor. However, others argued that the enactment of the Road Safety Acts (which mandates to lower BAC limits and enforce the use of seat belts) as well as the enactment of other vehicle safety standards have effectively reduced the number of fatalities (Votey, 1984). Oppe (1989) explored the reason behind this decrease since the explanations given in previous studies were not

really satisfactory. He proposed a new mathematical formula which shows that the number of fatalities is the product of vehicle-kilometres and rate of fatalities.

According to the Institute for Road Safety Research (SWOV), the decrease is supposed to be a combination of all the efforts made to improve the traffic system, which includes improvements in road systems, vehicle designs, crash measures, legislations, education and individual learning. These efforts are referred to as the 'learning curve for the society' (Oppe, 1989). However, Gaudry and Lapparent (2013a) perceived that the turnaround of the fatalities which occurred in 1972–1973 is the 'mystery of 1972–1973'. This is due to the lack of analyses on the single most important feature of road safety since the 1950s.

2.6 Road safety measure and intervention

Various interventions and measures have been implemented by countries over the years in order to improve their road safety. The enactment of road safety legislations has provided numerous opportunities for researchers to investigate the impact of legislations on the number of road traffic casualties. There is a large number of studies related to this subject in the literature. One of the earliest studies in this area was carried out by Jokschi and Wuerdemann (1973), whose objective is to forecast the decrease in fatalities until 1980. According to Haight (1994), the measures primarily taken to facilitate mobility are often justified by referring to their effect on road safety. Some of the important examples of these measures are: (1) driver education, (2) traffic rules and their enforcement, and (3) engineering of roadways. In promoting the 3E's, safety is typically the most persuasive argument due to the fact that failures in road safety may result in heart-wrenching consequences compared to failures in mobility. Baguley (2001) proposed the addition of a fourth E (Encouragement) to the current 3E's (Education, Enforcement and Engineering).

Bossche et al. (2004) investigated whether the number of accidents and victims in Belgium are influenced by weather conditions, economic conditions and policy regulations. They found that laws and regulations have a significant effect on road safety besides weather conditions. However, the effect of economic conditions on road safety seems to be negligible. Hermans et al. (2005) established that the laws concerning child restraints are the variable which significantly influence the number of fatalities on motorways. Moreover, Hermans et al. (2006) investigated the monthly frequency and severity of road traffic accidents in Belgium within 1974–1999. They found that the implementation of laws concerning seat belts, vehicle speed and alcohol consumption are effective in reducing the severity of road traffic accidents. In contrast, economic conditions have a negligible impact. Kim et al. (2006) evaluated the effects of five traffic safety policies implemented in Korea. They perceived that the following four interventions will help in reducing the number of fatalities: (1) mandatory seat belt law for front seats, (2) three-strikeout law for driving while intoxicated (DWI) offences, (3) heavier penalties for DWI offences, and (4) traffic violator reporting reward programme.

Gargett et al. (2011) highlighted that the following road safety initiatives are effective: (1) the introduction of compulsory seat belt use, (2) reduction in speed limits and the implementation of vehicle speed-related measures such as speed cameras, (3) random breath testing, (4) campaigns which promote awareness on the dangers of drunk driving and (5) programmes which are targeted at locations with high rates of crashes (e.g. the black spot programme). Another approach that has been used increasingly in recent years is the implementation of national road safety strategies that include specific targets for improvements in road safety outcomes. Zhang et al. (2011) analysed the road safety trends in China from 1951 to 2008. One explanation for the declining trend is the improvement in road traffic control, engineering, as well as other traffic safety measures implemented in China. Since 2004, many safety interventions have been made to curb road traffic

injuries (RTIs), which include: (1) the enforcement of speed limits, (2) the use of speed radar systems, (3) the provision of median barriers, and (4) traffic safety awareness campaigns.

Weijermars and Wesemann (2013) highlighted that numerous measures have been taken in the Netherlands since the 1970s concerning: (1) traffic legislation, (2) protection of car occupants, (3) motorcyclists and moped drivers, and (4) upgrading the safety features of infrastructure. Antoniou and Yannis (2013) analysed the macroscopic road safety trends in Greece. The interventions incorporated in their models are: (1) the 1986 financial crisis, (2) the 1991 old-car exchange scheme, and (3) the 1996 new road fatality definition. They found that all of the interventions are statistically significant.

According to Oster and Strong (2013), there are multiple causes of road traffic accidents and there can be interactions between factors such as restrained use, alcohol consumption, speeding, fatigue and distractions. Such interactions are not well understood and may be important. In addition, there are interventions which can affect travel patterns, which make evaluations difficult. Traffic safety is influenced by three broad sets of factors: (1) road and traffic environment, and engineering, (2) vehicle characteristics and performance, and (3) driver behaviour and performance. Major improvements have been made in both highway and vehicle design, as well as in crash countermeasures and mitigation.

In addition, Evans (2004) (cited by Oster and Strong (2013) stated that even though most studies indicate that the three sets of factors mentioned above are important, it is the third set of factors (i.e. driver behaviour and performance) that is most significant. Road and traffic engineering initiatives are perceived as having produced larger benefits in terms of risk reduction compared to changes in vehicles. Even though public policies certainly play a key role in decreasing the risk of road traffic accidents, it is evident that

changes in travel patterns and demographics are the important factors which will help reduce highway fatalities.

2.6.1 Seatbelt and other motor vehicle safety standards

Since it has been shown that the use of seat belts helps reduce the severity of road traffic accidents, the authorities have imposed that road users need to comply with motor vehicle safety standards. In the late 1960s, several countries comprising the US, the UK, France, Australia, Japan, Sweden and Germany made it compulsory for road users to fasten their seat belts. In the early 1970s, the use of seat belt was made compulsory for drivers. The National Traffic and Motor Vehicle Safety Act of 1966 was passed into a law in the US. Among the many aspects of the 'initial standards' are the requirement that all new cars manufactured for sale in the US must have outside rear view mirrors, side marker lights, redundant braking systems, energy-absorbing steering assemblies, high-penetration resistant windshields, more crash-resistant door locks, shoulder harnesses in front of outboard seats, and lap seat belts for all seating positions (Robertson 1977).

The results of study carried out by Joksch and Wuerdemann (1973) showed that the number of fatalities due to motor vehicle accidents in the US will decrease from 68,000 to 31,000 deaths in 1980, provided that crash countermeasures are installed in all vehicles and all of the car occupants fasten their seat belts. Scuffham and Langley (2002) discovered that there is a substantial reduction in the number of fatal crashes which is associated with the oil crisis in 1979 and seat belt wearing laws. The 1984 universal seat belt wearing law is associated with a sustained reduction in fatal crashes of 15.6%. These road policy factors appear to have a greater influence on crashes compared to the role of demographic and economic factors. The 1% increase in the open road speed limit is associated with a 0.5% increase in fatal crashes in the long term.

Beenstock and Gafni (2000) suggested that a decrease in the rate of road traffic accidents reflects the advancements of road safety technologies integrated in motor vehicle and road designs, rather than the effectiveness of road safety policies. Noland (2004) concluded that the improvements in fuel efficiency do not affect traffic-related fatalities and downsizing SUVs is probably a desirable policy to reduce both the fuel consumption of vehicles and the number of road traffic fatalities. Broughton and Knowles (2010) examined the national target that was set to reduce the number of road accident casualties in Great Britain. The results showed that improvements in secondary features of cars over the past 15 years have the most significant effect on reducing the total number of casualties. Akloweg et al. (2011) determined the relationship between road traffic accidents and servicing of used cars. The findings showed that there is a strong correlation between vehicle age and road traffic accidents. This is particularly true since old car models do not include safety features which will help minimize the impact of road traffic accidents.

2.6.2 Maximum blood alcohol concentration

Since the late 1960s, a number of countries such as the UK, Canada, Japan, the Netherlands and the US have introduced the legal limit for road users driving under the influence of alcohol by prescribing the maximum blood alcohol concentration (BAC). It is an offence to drive a motor vehicle with a BAC above the maximum legal limit. Ross (1973) evaluated the impact of implementing the Road Safety Act of 1967 in Great Britain. This Act consists of scientific tests used to determine and define the crime of drunk driving. This study provides evidence that subjective certainty of punishment can deter socially harmful behaviour such as drunk driving in Great Britain.

2.6.3 Speed limit

In 1964–1965, a relatively high speed limit (i.e. 70 mph) was introduced in the UK (see DfT, 2009). The US set a speed limit of 55 mph in 1974. Haney and Weber (1974) reviewed a large number of studies published since the 1960s in order to determine the relationship between vehicle speeds and motor vehicle accidents. This exhaustive review shows that lowering the speed limits will decrease: (1) the number of vehicles travelling at high speeds, (2) the mean speed of traffic, (3) the dispersion of speeds about the mean, and (4) the number of serious-injury and fatal-injury accidents.

One of the significant measures is the legislative changes that take place in Denmark, in the form of reduced speed limits. In 1979, the speed limits on rural roads and freeways were lowered from 90 to 80 km/h and from 110 to 100 km/h, respectively. However, in 1985, the urban speed limit in Denmark was lowered from 60 to 50 km/h. The first speed limit reduction appears to have a significant effect on road casualties, reducing the rate of fatal accidents by 21%, whereas the second speed limit reduction is statistically insignificant (Fridstrøm et al., 1995). Vernon et al. (2004) analysed the effects of increased speed limits on the rate of crashes, rate of fatalities and rate of injuries on motorways in Utah. The results showed that the increase in speed limits has an adverse effect on the occurrences of crashes for different types of crashes and motorways – however, the effect is not pronounced.

Lave (1985) came up with a different conclusion on the effect of speed limits on the number of road casualties. It has been widely accepted that speed is a factor that results in an increase in the number of casualties – however, this belief has been refuted in this study. According to Lave (1985), it is the variability in vehicle speed that kills, rather than speed itself. Koornstra (2007) highlighted that speed is not the primary factor that kills since traffic calming areas and motorways are the areas with the lowest fatality risks. The

speed limit in traffic calming areas is 30 km/h speed limit, whereas the speed limit in motorways is 100 or 120 km/h in the Netherlands. However, there are no speed limits for motorways in Germany. The fatality risk is highest where speed limits are inappropriately high (80 or 100 km/h speed limits on roads with mixed traffic, level crossings, and opposing traffic without mid-barriers).

According to the ONISR (2001), three out of five drivers who are involved in road traffic accidents violate the speed limits on rural roads in France (cf. Letirand & Delhomme, 2005). Zhang et al. (2011) reported that the decreasing rate of traffic crash deaths per 10,000 vehicles or per 100,000 people in China may be attributed to slower travel speeds resulting from an increase in traffic congestion as well as more experienced drivers.

2.6.4 Speed Limit Enforcement Camera

Feng (2001) reviewed the factors which affect road safety (with focus on speed limits and the use of speed cameras), and concluded that road safety is closely related to speed. However, the findings are rather inconclusive regarding the effectiveness of speed cameras in improving road safety. Hess and Polak (2003) investigated the effects of Speed Limit Enforcement Cameras (SLECs) and the results showed that there is an average decrease in the monthly accident frequency by around 18%. A more detailed analysis suggested that the introduction of SLECs decreases the percentage of road traffic injury accidents by 31.26%, which clearly shows that the implementation SLECs significantly decreases the number of accidents. The impact of speed management schemes was evaluated by Mountain (2005) who found that engineering schemes which incorporate vertical deflections (e.g. speed bumps or cushions) offer the largest benefit: an average

reduction of personal injury accidents of 44%. However, the reduction in personal injury accidents is 22% at sites where safety cameras were used to monitor speed limits.

According to a survey by Soole (2009), the subjects perceived that the installation of fixed cameras is most effective on freeways (69.7%), school zones (57.5%) and crash black spots (53.0%). However, the subjects perceived that marked patrol vehicles parked on the side of the road is most effective on freeways (62.7%) and school zones (56.3%), but only somewhat effective on urban roads (49.4%). However, others perceived that unmarked patrol vehicles in the traffic flow is most effective on freeways (57.0%). According to De Pauw et al. (2014), speeding remains a major challenge in road traffic safety. They examined the evolution in the number of crashes at various locations in which speed cameras are installed. The results showed that there is an insignificant decrease in the number of injury crashes, with a value of 8%. The number of severe crashes with serious and fatal injuries decreases by 29%. It can be concluded that speed cameras have a favourable effect on traffic safety, mainly on the severity of crashes.

2.6.5 Provision for motorcyclist

Evans and Frick (1988) found that the use of crash helmets reduces the risk of fatalities on motorways by roughly 27%. Stolzenberg and D'alessio (2003) analysed the repeal of the mandatory motorcycle helmet law for operators and passengers above 21 years old in Florida, US. The results showed that the repeal of mandatory motorcycle helmet use neither increases the rate of fatalities nor the rate of serious injuries due to road traffic accidents. Muller (2004) examined the law that exempts adult motorcyclists and moped riders from wearing helmets in Florida. The exemption is only applicable for those who have a medical insurance of \$10,000. The findings of the study indicate that this law should be revoked since the passing of this law actually increases the number of

motorcyclist fatalities. The number of motorcyclist fatalities is the highest, with a value of 48.6%.

Al Bavon and Standerfer (2010) sought to determine the effect of the implementation of the motorcycle helmet law in Texas on the rate of fatalities since the repeal of the universal helmet law in 1997. The results showed that there is a drastic increase in the rate of fatalities following the implementation of the law. The rate of fatalities, the rate of fatalities per registered motorcycle and the rate of fatalities per vehicle-mile travelled increases by 30, 15.2 and 25%, respectively after the repeal.

Radin Umar et al. (2000) analysed the impact of exclusive motorcycle lanes on the number of motorcyclist accidents along Federal Highway Route 2 in Malaysia. The results showed that the number of motorcyclist accidents is directly proportional to the cubic power of traffic flow and the number of motorcyclist accidents is reduced by approximately 39% with the existence of motorcycle lanes. Law et al. (2005) examined the effects of motorcycle safety intervention (MSP) and the concurrent economic crisis on the number of motorcycle-related accidents, injuries and fatalities in Malaysia. They found that the MSP is effective in reducing the number of motorcycle-related accidents, casualties and fatalities, with a value of 25, 27 and 38%, respectively.

Muhammad Marizwan et al. (2013) developed a model to forecast motorcyclist fatalities on various main roads in Malaysia. The results showed that the motorcycle fatalities per kilometre is influenced by the average daily number of motorcycles and the number of access points per kilometre. These parameters are statistically significant. In addition, the model estimates indicate that there is an increase in the number of access points per kilometre and average traffic volume (average daily number of motorcycles and average daily traffic volume), which are highly associated with the increase in motorcyclist fatalities per kilometre.

2.6.6 Driver licensing

Hedlund et al. (2003) reviewed literature concerning graduated driver licensing laws and they concluded that these laws reduce both road traffic fatalities and involvement in crashes among adolescents by increasing the minimum age at which the adolescents are allowed to drive as well as improving their driving skills. Izquierdo et al. (2011) examined the contribution of the penalty point system (which is the most important legislative measure for driving licences) in reducing the number of fatalities on roads in Spain over a period of 24 hours. The results showed that the introduction of the penalty point system has a very positive effect in reducing the number of road traffic fatalities and this effect remains positive to date. This success may be due to the continuing increase in surveillance measures and fines as well the growing interest in road safety publicized by news media since the measures were introduced. This leads to positive changes in driving behaviour.

2.6.7 Laws enforcement

Koornstra (2007) identified four types of behaviour that increases the risk of fatalities: (1) refusal to wear seat belts for car occupants, (2) refusal to wear safety helmets especially for riders of motorized two-wheelers, (3) drunk driving and (4) violations of speed limits. Road safety can be improved significantly by intensive enforcement of helmet usage (for motorcyclists and moped riders) and seat belt usage (for drivers and both front and back seat passengers), and applying stringent laws on drunk driving and speed limits. Improvements in passive vehicle safety has probably saved about 20% fatalities in the last two decades. Yannis et al. (2007) found that there is a clear relationship between the level of police enforcement and the reduction of traffic accident casualties.

Andrade et al. (2008) examined some measures that are in practice in Brazil and found that they have a small effect on the mortality of victims of road traffic injuries such that the rate of mortality exceeds 35 per 100,000 population. These measures comprise: (1) the obligatory use of seat belts in urban areas, (2) the implementation of speed control radar at several strategic points, and (3) the introduction of pre-hospital attention for victims of road traffic events. Lateef (2011) assessed the patterns of road traffic injuries (RTIs) and fatalities in Karachi. They made the following recommendations in order to reduce RTIs and fatalities: (1) enforce road safety regulations, (2) provide training programmes for motorcyclists, and (3) develop pedestrian-friendly infrastructures.

2.7 Road safety policy and target

Yannis et al. (2011a) highlighted that analysis of past road safety patterns in developed countries provides insight into the underlying process that relates the motorization level with personal risk. This analysis is beneficial to predict the evolution of road safety in developing countries which may have not yet reached the same breakpoints. According to Oster and Strong (2013), public policies play an important role in reducing the overall highway fatality rates. However, it is challenging to determine the effectiveness of a particular policy in improving safety.

Lu (2007) highlighted that the annual number of fatalities in the Netherlands decreases substantially by 7.2% during the period of implementation of Phase I of the Dutch road infrastructure redesign programme (1998–2002). However, even though Phase II was not implemented from 2002 to 2006, the annual number of fatalities decreases dramatically by 23.9%. SWOV attempted to identify the explanatory factors which result in the abrupt decrease of fatalities in 2004 – however, none of the explanations are convincing. According to Elvik (2010a), in most cases, the actual number of people killed or injured

in road accidents exceeds the figures predicted according to a road safety programme. Road safety is influenced by a variety of factors and therefore, it is not possible for a road safety programme to address all of the issues concerning road safety. Hence, road safety can only be controlled to a certain extent by the implementation of road safety programmes. However, the effects of implementing a road safety programme on road safety achievement are unclear.

2.7.1 Sustainable safety vision

Sustainable safety vision is an approach used to improve the road safety scenario in the Netherlands in the last two decades. According to Wegman et al. (2008), there are five core principles in sustainable safety vision, namely: (1) functionality of roads, (2) homogeneity of masses and/or speed and direction, (3) predictability of road course and road user behaviour by recognizable road designs, (4) forgivingness of both the environment and road users, and (5) state of awareness by the road users. In order to achieve sustainable safety vision, a large number of proposals are drafted which encompass a wide range of road safety aspects such as infrastructure, vehicles, intelligent transport systems, education, regulations and their enforcement, speed management, driving while under the influence of alcohol and drugs, young and novice drivers, cyclists and pedestrians, motorized two-wheelers, and heavy goods vehicles.

Mohan (2004) reviewed studies pertaining to road safety development. Some of the key findings in this review are: automatic speed enforcement by means of speed cameras seems to be effective in improving road safety, the use of helmets is the single most effective safety measure available for motorcyclists since it reduces the probability of injuries and fatalities, there is no clear evidence on the relationship between knowledge and attitudes, as well as the relationship between knowledge and behaviour. Moreover,

the successful strategies are those which involve a combination of education and other approaches such as legislations, regulations and eliminating the barriers that hinder the implementation of road safety measures, the only effective way to encourage the use of seat belts and child seats among motorists is the enforcement of laws which make their use compulsory, education on pedestrian safety can improve childrens knowledge on road safety and induce changes in road crossing behaviour.

Luoma and Sivak (2013) made the following recommendations in order to improve the road safety scenario in the US in order to be at par with Sweden, the UK and the Netherlands: (1) reduce the BAC limit to 0.5 g/l for all states and introduce effective random breath testing, (2) re-examine the current speed limit policies and improve speed enforcement, (3) implement primary seat-belt-wearing laws in each state that are applicable to both front and rear seat occupants, and reward vehicle manufacturers for installing advanced seat-belt reminders, (4) reconsider road safety targets to focus on reducing fatalities rather than the rate of fatalities per distance driven, and (5) develop new strategies to reduce the vehicle distance driven.

2.7.2 Road safety target

2.7.2.1 Setting road safety target

Elvik (2010a) highlighted that the road safety programmes implemented in Denmark, Finland and Sweden do not produce the desired effect as initially hoped. Wittenberg et al. (2013) concluded that the 50% reduction target set by the EU is too ambitious, considering that the total number of people killed in road accidents in 2001 is 54,000. Even though the number of road accident victims has decreased considerably, achieving the maximum permitted limit of 27,000 deaths in 2010 is overoptimistic. Improving the road safety scenario in the EU is possible provided that all of its Member States contribute

significantly to attain the collective objective of reducing the number of traffic road fatalities by half. These include revising the speed limits, making amendments to the seat belt laws, imposing stricter penalties for drunk driving, constructing safer roads, and producing safer vehicles. More importantly, each Member State should identify which measures are most effective in its country. Moreover, Lassarre (2001) claimed that Europe's road systems are capable of absorbing a 6% increase in traffic (vehicle-km) per annum while maintaining a constant number of fatalities.

Elvik (1993) examined the safety performance of Norwegian countries in which quantified road safety targets were set, and the best performance was achieved by countries with highly ambitious quantified targets. The more recent studies by Wong et al. (2006) and Allsop et al. (2011) confirmed Elvik's conclusion. According to Bener (2009), the prevention of road traffic injuries in developing nations is inhibited by the lack of knowledge, absence of reliable estimates on the current level of injuries, and restrictive views on health and diseases. This situation occurs because studies that are focused on reducing road traffic injuries are funded at a disproportionately lower level compared to other health issues. In addition, the governments in developing countries do not perceive road traffic injuries and fatalities as a public health problem.

2.7.2.2 Vision zero

According to Kanellaidis and Vardaki (2011), Vision Zero in Sweden, Sustainable Safety Vision in the Netherlands, and Australia's safe system are the main safe-system strategies in the world. They highlighted the role of road safety auditing to improve road designs and develop a safety culture among road designers. Corben et al. (2010) described the approach used in Western Australia (WA) in support of the 'Towards Zero' strategy. The 'Safe System Matrix' concept was created to identify the best mix of

initiatives for Western Australia's next road safety strategy. The matrix consists of the following key areas: (1) safe roads and roadsides, (2) safe speeds, (3) safe vehicles, and (4) safe road use.

Gaudry and Lapparent (2013b) highlighted the most interesting aspect of Annex (ITFOCE/JTRC, 2008b: 439). The objective of reducing the death rate due to road traffic accidents through the establishment of national targets is highlighted in this report. It shall be noted that national targets are not forecasts – rather, they are extensions of the current trends. Amazingly, it is observed that these targets are all linear with a decreasing trend, except for three jurisdictions out of the 40 jurisdictions investigated. They also observed a few trend breaks, which indicates that the national targets are achieved only if the trend is downwards – literally, upward trends are non-existent. Gaudry and Lapparent (2013b) also highlighted the following statement given in the official brochure of the Swedish Road Administration: 'Since the adoption of Zero Vision, the death toll on Swedish roads has declined'. This implies that the objective of the Zero Vision has not progressed significantly since its adoption.

2.8 Recent road safety reports by WHO

In 2004, Peden et al. prepared a report entitled '*World Report on Road Traffic Injury Prevention*' which was issued under the WHO and World Bank. It is highlighted in the report that even though the decrease in road traffic fatalities is forecasted to be around 20% in high-income nations, the current and forecasted trends reflected by both low-income and medium-income nations indicate that there will a large increase in global road crash mortality between 2000 and 2020. The majority of road traffic accident victims in low-income countries are pedestrians, passengers, cyclists, users of motorized two-

wheelers, as well as occupants of buses and minibuses. In contrast, the majority of road traffic accident victims in most high-income countries are car occupants.

Peden et al. (2004) also reported that road traffic injuries are the eleventh leading cause of deaths in 2002. In economic terms, the cost of road crash injuries is estimated to be roughly 1, 1.5 and 2% of the gross national product (GNP) for low-income, middle-income and high-income countries, respectively. It is also highlighted in the report that human errors are the leading factor in road traffic accidents. Males account for 73% of all road traffic fatalities whereby the overall rate of fatalities is nearly thrice the rate of fatalities for females. According to the report, commuting by buses and trains is generally safer compared to other modes of travel. The rate of fatalities per vehicle-kilometre travelled is deemed to be a useful parameter for comparing road safety achievement between different countries – however, this parameter does not account for non-motorized travel. High vehicle speeds beyond the maximum speed limits are common and may contribute to 30% of road traffic accidents and fatalities. Automatic speed enforcement by means of speed cameras is now employed in many countries. Speed cameras are used to provide visual evidence of a speeding offence that is admissible in court and has been proven to be a highly effective means of speed enforcement in high-income countries.

In 2009, the WHO released a report entitled '*Global Status Report on Road Safety*'. The report is the first broad assessment of the status of road safety in 178 countries. According to the report, over 1.2 million people die each year on the world's roads, whereas 20 to 50 million people suffer from non-fatal injuries due to road traffic accidents. The rate of road traffic fatalities is higher in low-income and middle-income countries, with a value of 21.5 and 19.5 per 100,000 population, respectively. In contrast, the rate of road traffic fatalities is only 10.3 per 100,000 population in high-income countries. It is shown that over 90% of the world's fatalities on the roads occur in low-

income and middle-income countries, considering that these countries only make up 48% of the world's registered vehicles. There is a declining trend in the rate of fatalities since the last four to five decades in most high-income countries. In addition, road traffic injuries remain an important cause of fatalities, injuries and disabilities, even in high-income nations (WHO, 2009).

2.9 Risk compensation

In general, it is widely accepted that motor vehicle safety provisions will effectively reduce road traffic fatalities. However, Peltzman (1975), who examined the impact of the National Traffic and Motor Vehicle Safety Act of 1966 on the rate of road traffic fatalities, opposes this point of view. He concluded that safety regulations have no effect whatsoever on highway death tolls. In addition, he hypothesized that the National Traffic and Motor Vehicle Safety Act of 1966 will actually increase driving intensity such as speeding, which will endanger vulnerable road users. Peltzman's findings have received criticisms ever since they have been published.

2.9.1 Researchers who opposed the risk compensation theory

Among the earlier studies which oppose Peltzman's views are the works of Joksche (1976) and Robertson (1977), and they concluded that the implementation of motor vehicle safety standards significantly reduces the number of road traffic casualties. Other studies also showed contradictory results. The NHTSA (1984) revealed that the use of lap/shoulder belts reduces fatalities by 40–50%. Evans (1986) discovered that the fatalities of drivers and front passengers in the US can be reduced by 43% if all of the present unbelted drivers and front passengers used lap/shoulder belts without changing their behaviour. Wagenaar et al. (1988) examined the effects of compulsory seat belt use

on the number of occupants fatally injured in traffic crashes. The results revealed a statistically significant decline (8.7%) in the rate of front-seat passenger fatalities. They also presented various studies on the reduction in the rate of front-seat passenger fatalities due to compulsory seat belt use. The estimated reduction in fatalities following seat belt use varies from one country to another, ranging from 0 to 80%.

Peltzman (1975) conclusion that regulation of motor vehicle safety standards will increase driving intensity (reckless behaviour) has also been greatly debated in the literature. Since then, Peltzman's hypothesis has been termed as 'risk compensation', 'offsetting behaviour' or 'risk homeostasis' (Asch et al., 1991). The risk compensation theory was popularized by Wilde (1982) in one of his major works entitled '*A theory of risk homeostasis: Implications for safety and health*'. According to the theory, the improvement in safety per unit distance of mobility increases the drivers' moving speed and mobility per head of population – however, it has no effect on the annual rate of traffic accidents per capita. The risk compensation theory leads to the conclusion that the only measures which have a permanent effect on accidents are those that will alter the attitudes of drivers towards risk-taking (Hakim et al., 1991; Johansson, 1996). Peltzman (1975) and Crandall et al. (1986) (cited by Hakim et al. (1991)) conducted independent studies and the results showed that the National Traffic and Motor Vehicle Safety Act significantly reduces the rate of fatalities for car occupants. However, it is worth to note that the rate of fatalities for other road users (pedestrians, cyclists and motorcyclists) increases concurrently.

The risk compensation theory challenges the foundations of injury prevention strategies. According to this theory, the effective safety measures are those that will alter one's desired risk level, whereas others that merely modify the environment or regulate one's behaviour without affecting one's target risk level is basically useless (Hedlund,

2000). Robertson (1981) who conducted a study similar to Peltzman's also oppose this theory, and found that there is no evidence of risk compensation. He found that there is a decrease in road traffic fatalities following the implementation of federal safety regulations for pedestrians, cyclists, motorcyclists and car occupants. According to Preusser et al. (1988), the following drivers are less likely to wear seat belts: (1) young drivers, (2) drivers under the influence of alcohol, (3) drivers who run red lights and (4) drivers who follow preceding vehicles closely beyond a space that is considered reasonable and safe. In addition, the risk homeostasis theory posits that the number of traffic accidents per unit time of driving also tends to remain constant, and is essentially independent of changes in the traffic safety system (Evans, 1986).

Evans (1986) conducted a study to examine the validity of this claim using a wide variety of traffic accident data, and concluded that the risk homeostasis theory should be rejected because there is no convincing evidence to support the theory – rather, most of the evidence refutes the theory. Roh et al. (1999) applied Peltzman's formula over a longer time period (1973–1994) and found that the formula does not fit well with the data used in their study, producing unacceptable forecast errors. Nevertheless, the OECD (1990) (cited by Gossner and Picard (2005)) concluded that behavioural adaptation to changes in the road transport system exists and affects the safety benefits achieved through road safety programmes. The results showed that behavioural adaptation generally does not eliminate the safety gains obtained – rather, it reduces the effectiveness of road safety programmes in a number of cases. According to Graham (1982), risk compensation is more likely to occur for policies designed to reduce the frequency of accidents rather than those designed to reduce the severity of accidents. For instance, the use of motorcycle helmets will certainly reduce the severity of accidents, but not necessarily the frequency of accidents, whereas improvements in road safety will

probably reduce the frequency of accidents and hence, results in more compensating behaviour.

2.9.2 Researchers who agreed with risk compensation theory

Even though the majority of researchers disagree with Peltzman, there are others who support his view. Conybearre (1980), who evaluated the impact of automobile safety regulations in Australia, supported the findings of Peltzman. This study provides evidence that the use of seat belts and other road safety measures will increase the driving intensity of drivers. This has an undesirable consequence – the rate of road traffic casualties of motor vehicle occupants does not decrease as much as initially expected. The rate of road traffic casualties for motor vehicle occupants actually increases with the implementation of these measures. Adams (1982) also observed a similar trend regarding the effect of seat belt use on the number of fatalities. He conducted a survey on countries that constitute over 80% of the world's car population, and the results revealed that protecting car occupants from the consequences of bad driving actually encourages bad driving behaviour. The decrease in road accident fatalities is greater in countries which do not pass seat belt laws.

Some more recent studies also supported the idea that risk compensation theory occurs in road safety. According to Adams (1985), risk-taking behaviour involves the attempt to balance perceived risks and desired risks, and people generally adapt their behaviour in response to changes in perceived risks. He also refuted the idea that seat belt legislation saves a large number of lives which is frequently forwarded by advocates of public health measures. About a decade later, Adams (1994) criticized the work of Durbin and Harvey (1986), in which he concluded that none of the time series models developed by them consist of alcohol-related variables. According to Durbin and Harvey (1986), the decrease

in the number of fatalities below the projected trend in 1983 is attributed to the enforcement of the seat belt law and none is attributed to campaigns against drunk driving. They estimated that the percentage of fatalities for pedestrians and cyclists is 8 and 13%, respectively. In addition, they estimated that there is an increase in the percentage of fatalities for passengers in rear seats by 27%, considering that the law is inapplicable to these people. It is interesting to note that the number of pedestrians and cyclists killed by heavy vehicles and public service vehicles (road users which are not governed by the seat belt law) decreases following the enforcement of the law.

Simonet and Wilde (1997) presented the Munich taxicab experiment in support of the theory. Some of the taxis in a taxi fleet in Munich, Germany, were equipped with anti-lock brakes (ABS). The results showed that there is a significantly higher frequency of extreme deceleration (i.e. extremely hard braking) in taxis equipped with ABS. The results revealed that the drivers of ABS taxis made sharper turns in curves and they were less accurate in lane-holding. These taxi drivers also proceeded at a shorter forward sight distance, and frequently produced critical incidents such that other road users had to take action in order to prevent collision. These differences were found to be statistically significant – the ABS taxis were also observed to move significantly faster at one of the four measuring points. These results clearly demonstrate that drivers show a change of behaviour in response to the availability of ABS in the vehicle.

Assum et al. (1999) found that drivers exhibit risk compensation behaviour in the presence of road lights, which is reflected by speeding and reduced concentration. The results of the TAG-1 model developed by Jaeger and Lassarre (2000) in France showed there is a risk compensation effect since the use of seat belts leads to an increase in vehicle speed. This is known as driver behaviour retroaction (cf. Bossche & Wets, 2003). Hedlund (2000) deduced that risk compensation can occur since humans are not

machines. Hence, it can be expected that drivers will show a change in behaviour in response to changes in the environment. The implementation of road safety measures will indeed cause a change in the road environment, and thus leads to a change in the behaviour of drivers. There are a variety of rational and behavioural factors which will determine whether one's behaviour will change and the manner in which the behaviour is changed. It is rather simplistic for one to assume that a person's behaviour will never change.

In addition, according to Malnaca (2008), most researchers agree that people will adapt their behaviour to compensate for the risk that they perceive. This leads to the following question: How much do these people compensate? A couple of studies in recent years also indicate that risk compensation may occur and therefore, further analysis is required (Wang et al., 2013; Gaudry & Lapparent, 2013b).

2.10 Unit of exposure to risk

According to the ECMT (1984), the definition 'deaths per 100,000 cars' is not a sufficient criterion to be used as a basis for comparisons. In addition, the definition 'deaths per million population' should be used only for comparisons between countries with similar vehicle ownership ratios. However, it shall be noted that this definition does not necessarily indicate that a country's road safety policy is better or worse than another's (cf. Andreassen, 1991). Wegman and Oppe (2010) suggested that the use of mortality rates (the number of fatalities per head of the population) should be used to compare traffic risks – however, the use of mortality rates is negated by the fact that the motorization level is not taken into account. For this reason, the fatality risk is commonly used as a criterion to assess road traffic safety, and it is defined as the number of fatalities per motor vehicle kilometre. However, it shall be noted that if the data on motor vehicle

kilometres are unavailable, then the rate of fatalities should be used, which is defined as the number of fatalities per motor vehicle.

Baguley (2001) defined that it is very difficult to take the 'exposure to risk' into account since it can be expected that the average mileage of various road user groups may not necessarily be monitored by means of regular surveys. Even though the number of deaths per 10,000 registered vehicles is far from being a perfect measure, it is commonly used to assess and compare road traffic achievements between countries because these data are typically recorded in most countries. For instance, if the number of fatalities per 100,000 population is used, a country such as Bangladesh will appear to be 'safe' since the number of population is high but the number of vehicles is low. Likewise, a developing country such as Malaysia – a country which experiences tremendous traffic growth relative to its small population – will appear to perform extremely poor in comparison.

Hakkert and Braimaister (2002) detailed the definitions of accident, exposure and risk. The term 'accident' is defined as an event that occurs on a public road which results in injuries and at least one motor vehicle is involved. It is generally understood that the term 'exposure' refers to the exposure to risk, and it is defined as some form of the amount of travel, either by vehicle or on foot. In the field of road safety, 'risk' is used as a way to quantify the level of road safety relative to the amount of exposure, as opposed to the absolute level of safety as measured by the absolute number of accidents or casualties. However, the lack of comprehensive data on exposure is one of the reasons that comparisons are usually made based on per capita or per vehicle, and most of the data primarily deal with fatalities because of the inaccuracies in the reported number of injuries.

Shen et al. (2012) highlighted that the population size, the number of registered vehicles and the distance travelled are the measures frequently used to assess the exposure to risk. Nevertheless, there has been considerable debate in the past regarding which measure is most appropriate for use as an indicator of the exposure to risk since the definition of risk itself is subjective and is not consistent in most cases. Hence, Shen et al. (2012) implemented the data envelopment analysis (DEA) approach and its extensions to provide an overall picture of road safety risks in 27 Member States of the EU, as well as identify specific benchmarks and assign practical targets for the number of road fatalities in the underperforming states. They used three common measures as the inputs of the model to indicate the exposure to risk, namely the number of inhabitants, passenger kilometres travelled and passenger cars. The number of road fatalities was used as the output of the model.

In addition, according to Shen et al. (2012), the mortality rate an important criterion for road safety evaluation since it permits comparisons with other causes of death such as heart disease. The mortality rate is defined as the number of fatalities per million inhabitants. However, this risk indicator is undesirable for comparing traffic risks since it excludes the motorization level. For this reason, the risk of fatalities is introduced, which is an estimation of the exposure to risk in terms of the traffic volume. The risk of fatalities is defined as the number of fatalities per distance travelled – for example, the number of fatalities per 10 billion passenger-kilometres (pkm) travelled. This risk indicator is favoured by road transport authorities since it implicitly discounts the rate of fatalities if travel increases. However, the definition of the risk of fatalities differs significantly across countries, and only a few countries have gathered data on this measure of exposure. Hence, the rate of fatalities is typically used and it is defined as the number of fatalities per million registered vehicles. However, this risk indicator differs from other risk indicators in the sense that the annual distance travelled is unknown.

In addition, Antoniou et al. (2012) provided a detailed description on the role of macroscopic modelling in developing countries. For example, knowledge on the expected breakpoints in road safety fatality trends (obtained from the time series data in developed countries) can be applied to generate accurate predictions for developing countries. Understanding why fatality breakpoints as well as trend reversal occurs is crucial in order to improve road safety in countries which have not yet reached the motorization levels of developed countries. In addition, exposure data are very useful in road safety analysis since they help illustrate the underlying trends that lead to the present road safety scenario. The main exposure measures used are the vehicle kilometres travelled, the person kilometres travelled or the time travelled. However, gathering or estimating exposure data is much more challenging and these data are often unavailable for analysis. One way to overcome this limitation is to seek proxies (i.e. available data or data that can be easily gathered) that have high correlation with the actual exposure data. Some examples of proxies are the number of vehicles in circulation or the amount of fuel sold at gas stations. Time series methods have been used to account for and correct temporal correlations which are typical of macroscopic road safety data.

2.11 Road safety index

Al Haji (2005) proposed a set of methodologies which integrate different indicators of road safety into a single index. This index is known as the Road Safety Development Index (RSDI). In this study, the road safety scenario and trends were compared empirically between 10 Southeast Asian countries and Sweden from 1994 to 2003. The results showed that there is a remarkable difference in the road safety scenario between ASEAN countries, even between those with the same motorization levels. Singapore and Brunei appear to have the best RSDI record among the ASEAN countries based on the indicators used in the study. However, it was found that Laos, Cambodia and Vietnam

have low RSDI records. Malaysia is ranked fourth based on the RSDI method. However, based on the number of fatalities per 100,000 population, Malaysia ranks last among ASEAN countries in 2003.

Hakkert et al. (2007) developed safety performance indicators (SPIs) as a means to compare European countries by means of a road safety performance index. The SPIs are basically measures that reflect the operational conditions of the road traffic system, which influence the system's safety performance. As the name implies, these indicators help illustrate the effectiveness of a road safety programme. They help one to answer the following questions: Does the road safety programme fulfil its objectives? Is the desired outcome achieved with the implementation of the road safety programme? Seven problematic areas are selected during the development of SPIs: (1) alcohol and drug-use, (2) vehicle speed, (3) protection systems, (4) daytime running lights, (5) vehicles, (6) roads, and (7) trauma management.

2.12 Road safety modelling

Linear and non-linear models have been employed to forecast the trend of casualties in the future. The growth of knowledge on road safety in the 1970s have led to studies which merely use the explanatory variables and models available at the time rather than identifying new variables and developing new models. The lack of accuracy in model forecasts has also created doubts in extending research on the subject. These doubts were highlighted by Singleton (1984): *In all European countries there has been extensive growth in accident studies since about 1970. This was stimulated and was also made possible by the increasing affluence of the period. Most of this work was practical and pragmatic with an emphasis on the collection and collation of data from real situations. The inevitable result has been increased awareness of the industrial situation, but a*

dearth of new theoretical concepts. Most current theories are of the “systems” type which can explain anything but predict nothing.

According to Lateef (2010), road crashes are still seen as ‘accidents’ which cannot be prevented in most developing countries. It is vital that road traffic accidents are acknowledged as a major public health problem that can be predicted and prevented. Many high-income countries such as Australia have achieved dramatic success in reducing the incidence of road traffic injuries in recent decades despite the increase in motorization levels. Perhaps the strategies that have been proven to be effective in these countries can be used as guideline for developing countries to modify and implement intervention strategies which will reduce the occurrence of road traffic injuries.

2.12.1 Road safety modelling is complex

According to Votey (1986), modelling road accident and injury patterns is an arduous task. The first step involved in modelling road accidents is to formulate a conceptual framework of the problem, which incorporates all of the causal factors. In this step, the analyst needs to combine his or her theory, observations, knowledge and hypotheses in order to formulate a description of the accident generation process. Fridstrøm et al. (1995) highlighted that road accidents are unpredictable in nature – if they were predictable, they would not have happened. According to Anonymous (2010), road safety is rather complicated, and the design and evaluation of road safety programmes are not of interest to researchers at the time. In this study, a quick literature survey was conducted in April 2010 using SCOPUS and ISI Web of Science scientific literature online databases based on the search term ‘road safety management’, and the search results reveal that the number of articles that have been published on the subject since 1989 is less than 30.

However, the uncertainties concerning road safety management have motivated researchers to conduct research on road safety.

Antoniou et al. (2012) emphasized that modelling road safety is indeed complex, whereby it is necessary for the analyst to consider the quantifiable impact of specific parameters as well as the underlying trends which are not always measurable or observable. Among the examples of the parameters that cannot be modelled directly are: (1) the sensitivity of users to road safety campaigns, (2) the improvements in the quality of the vehicle fleet, (3) the improvements in the driving skills of the general population, and (4) the overall improvement in the conditions of road networks. Hence, the analyst needs to consider measurable parameters and time (which embodies all the remaining parameters) during modelling. Macroscopic modelling can provide insight into this problem and support policy-makers in developing countries to amend their policies in response to the changing macroscopic conditions. However, traffic fatality risk is also dependent on other parameters such as vehicle quality, traffic safety initiatives and regulations, as well as the intensity of police enforcement – and these parameters are not expected to influence the results of macroscopic analyses.

2.12.2 Road safety modelling is practical

Karlaftis and Tarko (1998) stated that road safety modelling attracts considerable research interest because of its wide variety of applications. More importantly, the WHO (2004) highlighted that road traffic crashes are predictable and therefore, they can be prevented. Hazen and Ehiri (2006) also agreed that most road traffic injuries can be prevented, and nowadays, there are significant improvements in knowledge and technology which will help in achieving this aim. According to Zhu (2010), even though the occurrence of traffic accidents is occasional, it can be predicted scientifically using

statistical indexes. Antoniou et al. (2012) stated that macroscopic road safety modelling and forecasting is an active research area, where ideas are being debated and interesting developments are still being made. It has been noted in the literature that advancements in road safety studies began in the 1980s (Hakim et al., 1991; Bergel-Hayat, 2012).

According to Hauer (2010), there are various ways to predict the number and rate of road traffic fatalities. Basically, for the same set of data, different prediction methods will produce different results. The targets set and the estimates of the effect of a road safety measure are generally dependent on the method used to predict the number and rate of road traffic fatalities. For this reason, it is crucial for the analyst to determine the best prediction method. Some methods rely on the extrapolation of past trends whereas others make use of comparison groups. Some methods are used to model the causal factors that govern the evolution of the safety of a unit. Some methods are suitable if the unit is an intersection or a road section whereas other methods are more suitable if the unit is a province, state or country. Some methods are suitable to make predictions in the next few years (short-term predictions) whereas other methods are more suitable to make predictions in the distant future (long-term predictions).

Elvik (2010b) examined the stability of the number of road traffic fatalities over the long term in eight highly motorized countries: Norway, Sweden, Denmark, Finland, Netherlands, Great Britain, Australia and the US. He observed that the decrease in road traffic fatalities is rather unstable over the years and more importantly, the trends that fit past trends very well are virtually useless for prediction purposes. Even though some models provide better fit for the data compared with simple trend lines, these models are not necessarily useful to predict future trends. This is because forecasting models require the prediction of future changes in all explanatory variables. It is important to note that past trends do not provide a reliable basis to predict the number of road traffic fatalities.

According to Broughton and Knowles (2010), forecasting the number of casualties at least 10 years in the future is an arduous task since the outcomes are dependent on the future behaviour of travellers and the growth of road travel as well as the means used to promote safe travel over a certain period. However, there is a method which has been proven to be successful in forecasting the number of casualties in Great Britain, as well as other countries. This method is based on the long-term relationship between the annual traffic volume in the country and the number of casualties. Several authors have observed that the rate of fatalities per hundred million vehicle kilometres travelled has declined over several decades at a nearly constant rate each year.

Gargett et al. (2011) stated that even though the method of using past observations to predict future outcomes is a standard forecasting method widely used among policy-makers, it is important for one to consider the limitations of the method. When forecasts are made based on past trends, it is inherently assumed that past behaviours, relationships and outcomes remain invariable over the forecasting horizon – however, this is not always the case when it comes to road traffic crashes.

2.12.3 Exponential model

Koornstra (1988) analysed the growth of vehicle kilometres for various countries over a long period and he deduced that the growth curves in the initial phase follows an increasing exponential trend such as logistic, Gompertz and log-reciprocal curves. Oppe (1989) worked further on the theory proposed by Koornstra (1988) based on the following assumptions: (1) there is a monotonically increasing S-shaped saturation curve with regards to the development of the number of vehicle kilometres per year, (2) there is a monotonically decreasing curve for the development of the rate of fatalities per year – this is known as the ‘risk curve’, and (3) the number of fatalities per year is determined

from these curves by multiplying their respective values. Gaudry and Lapparent (2013a) attempted to mathematically reproduce the turning point proposed by Oppe or Koornstra – however, they were unable to explain the trend because policies cannot be prescribed from non-structural equation curve fittings.

2.12.4 Extrapolation model

According to Hauer (2010), extrapolation is based on the assumption that time trends of the past will continue into the future. It is commonly believed that one will be able to produce a more accurate prediction, provided that the causes of the evolution of the process are well-understood and the magnitudes of these causes are known or at least, they can be predicted. It may be possible to describe how the expected number of accidents will change with changes in the values of some causal factors provided that the following conditions are fulfilled: (1) the data on all important causal factors are available, (2) the functional relationship between the expected number of accidents and causal factors is sufficiently understood, and (3) the difference between association and causation can be clearly distinguished in the regression model.

Moreover, in order to predict the number of accidents at a time point in the future, one must be able to predict the magnitudes of all causal factors at that time, which further complicates the prediction by causal models. For these reasons, it is not clear as to whether the predictions made based on data in statistical regression (structural) models are superior to those based on extrapolating time trends. In order to obtain a satisfactory explanation of the peak in the time series of fatalities, the analyst has to include a cause-effect representation of how risk (accidents per unit of exposure) changes as a function of causal variables (e.g. changes in public customs, investment in infrastructure, medicine, and urbanization) in the model.

In addition, at present, knowledge on these effects is rather inadequate. Even if the analyst knows how risk varies with respect to public customs or urbanization level, it is still necessary for the analyst to predict how this risk varies in the future. Extrapolation is still required unless the evolution of public customs or urbanization level can be described as a function of variables other than time. The hope that the evolution of fatalities can be explained by cause and effect without resorting to extrapolation of time leads to regress. Hence, for these reasons, only time-based extrapolation methods will be examined. The analyst needs to have some sort of a benchmark on the prediction quality in order to determine which of these two methods is better. The average bias and standard error of the average bias are used for this purpose (Hauer, 2010).

2.12.5 Structural time series model

Since the 1980s, in Canada, a specific class of accident models has been developed which includes a variety of variables. These models are classified under the common name 'DRAG' (Gaudry, 2002). DRAG modelling consists of three independently modelled stages: (1) traffic volume, (2) risk which is expressed in terms of accidents per kilometre and (3) the number of victims per accident. DRAG models are structural explanatory models, which consist of a relatively large number of explanatory variables, and whose partial effects on the exposure, frequency and severity of accidents are estimated by means of ARIMA(X) techniques (cf. Hermans et al., 2005). According to Bijleveld et al. (2008), the DRAG framework developed by Gaudry and Lassarre (2000) has been widely used in road safety research. However, the dependency of the DRAG model on economic and social variables such as unemployment and alcohol consumption (which are difficult to forecast) somewhat limits its application in forecasting.

Harvey and Durbin (1986) were apparently the first to apply structural time series analysis in road safety research. They forewent the Box and Tiao (1975) intervention analysis (which is based on Box and Jenkins (1970) ARIMA modelling) unlike what most researchers did at the time. They analysed the effect of compulsory seat belt use on the fatalities of car occupants and the results revealed that the number of fatalities reduce by 18 and 25% for car drivers and front seat passengers, respectively. Reurings and Commandeur (2007) reviewed studies carried out in countries other than the Netherlands which are focused on developing explanatory models to analyse and forecast developments in road safety. All of the studies used some form of disaggregation in terms of the severity of accidents. A number of studies disaggregated the traffic process according to the type of road user, type of road, gender, and/or age group. Classical linear regression and log-linear regression analyses were used in studies performed in Great Britain, whereas ARIMA and DRAG modelling were used in studies performed in Belgium, Canada, France and Sweden.

Both ARIMA and DRAG are dedicated time series models which account for dependencies between the observations in the time series data (Reurings & Commandeur, 2007). According to Broughton and Knowles (2010), various approaches have been developed to evaluate the developments in road traffic safety in different countries, and some of these methods have been used for forecasting. The ARIMA and DRAG2 models have been used in Belgium, Canada, France and Sweden. However, structural time series models are the preferred models in the Netherlands. Bijleveld et al. (2008) also highlighted that structural time series models are preferable from a methodological perspective. Moreover, the WHO (2009, 2013) used the negative binomial regression model to estimate global road traffic deaths.

Bergel-Hayat (2012) concluded in his review that temporal monitoring is no longer restricted to a straightforward conventional descriptive approach since the early 1980s. Rather, the purpose of temporal monitoring is to quantify the effect of explanatory factors on the risk and severity of accidents, as well as evaluate the effectiveness of road safety interventions. For this reason, explanatory models (i.e. models that include explanatory variables) have superseded descriptive models. Explanatory models make use of rich economic formulations along with sophisticated econometric specifications. In general, researchers will refer to a methodological framework in order to quantify the influence of factors related to the transportation system, travel demand, economy and road safety management. There are two types of risk factors: (1) internal factors (factors that are associated with the transportation system such as vehicles, drivers and infrastructure) and (2) external factors (factors that are associated with the environment such as climatic, economic, demographic and legislative factors).

Yannis et al. (2011a) developed parsimonious models which link the motorization level with the decrease in rate of fatalities across countries in the EU in the past three decades. They observed that the autoregressive log-transformed model outperforms the base autoregressive non-linear model, and enables them to identify the countries with the best and worst road safety performance. Yannis et al. (2011a) also highlighted that time series methods have been used to remove temporal trends and autocorrelations from traffic fatality risk models, which enables one to assess the impact of macroscopic road safety-related model parameters on traffic risk. According to Bergel-Hayat (2012), autoregressive models as well as those that are similar to autoregressive models once the observed data have become stable are the models commonly used for structural modelling. The most advanced explanatory autoregressive model for road risks was developed by Gaudry in 1984. According to Sukhai et al. (2011), NARX models are similar to ARIMA models that have been used extensively in road safety research. Such

models do not only contain autoregressive terms of the variable being studied, but also exogenous input terms that can be either autoregressive or obtained from the present time step.

2.12.6 Time series analysis and intervention model

Commandeur et al. (2013) described that the data collected to build a road safety observatory typically include observations that are made sequentially through time. Examples of such data (known as time series data) include the annual or monthly number of road traffic accidents, traffic fatalities or vehicle kilometres driven in a country, as well as the corresponding values of safety performance indicators (e.g. data on speeding, seat belt use and alcohol use). The assumptions inherent in some common statistical techniques are often violated due to the unique characteristics of time series data – one of them is serial dependency among disturbances associated with the observations.

The progress of time series analysis techniques as well as the availability of more detailed time series data has led to advances in statistical approaches in road traffic safety (Bijleveld et al., 2006). An important class of road safety models is based on time series analysis. A time series basically consists of repeated measurements of a certain phenomenon at regular intervals, and therefore, the evolution of road safety with respect to time can be studied in this manner (Hermans et al., 2005). Time series can be defined as a stochastic process in which the subsequent observations randomly change with time. The observations may be the number of people killed or the number of people injured in road traffic accidents, or the total number of accidents over a month, quarter or year. These data are used to create the time series which are then used to build the model (Krystek & Żukowska, 2007). ARIMA intervention analysis has been used in numerous studies pertaining to the assessment of traffic safety policies (Kim et al., 2006). However,

Box-Jenkins methods are more widely used for time series analyses compared to state-space methods (Hermans et al., 2006).

According to Bergel-Hayat (2012), decision-makers at the national level typically use time series analysis as a tool to describe and predict changes in road safety trends. One of the challenges for decision-makers is to identify the explanatory factors for changes in past trends, and intervene using appropriate measures in order to alter the trends in the future. For this reason, researchers have developed various models to provide the answers to these questions. Time series models have been applied on road safety indicators to explain past trends and predict trends in the future. In addition to the traffic volume or size of the vehicle fleet, risk exposure is also assessed with reference to economic activities (e.g. household income), price of fuel, structure of the population (e.g. proportion of young drivers), and regulatory measures (e.g. speed limits, the introduction of technical vehicle checks or the minimum legal age for alcohol consumption). The development of a road safety observation in Europe is limited by the lack of data, rather than the lack of appropriate methods.

The EC (2004) released the COST Action 329 report which includes a description on various models for traffic and safety development as well as interventions. One of conclusions made in the report is that there is strong potential in developing time series models for monitoring purposes and long-term forecasting on a broader scale. Hermans et al. (2006) suggested that an important class of road safety models is based on time series analysis. Box-Jenkins methods are more commonly used for time series analysis compared to state-space methods.

Antoniou and Yannis (2013) analysed the macroscopic road safety trends in Greece. Since reliable estimates of vehicle-kilometres were not available, the number of vehicles in circulation was used as a proxy to measure exposure. One of the modelling approaches

that have been frequently used to model a control process as an intervention is the interrupted time series technique. This technique is used to evaluate changes in the accident levels. More sophisticated versions of interrupted time series technique involve the use of univariate ARIMA models (Votey, 1986). In general, time series models are better than cross-sectional models because cross-sectional models do not account for geographical and cultural differences between countries, states or provinces (Page, 2001). Econometric models such as regression and time series models are very useful to enhance one's understanding on traffic safety trends (Bossche et al., 2004).

According to Antoniou and Yannis (2013), it is recommended that time series methods such as ARMA-type models (stationary data) and ARIMA or state-space models (non-stationary data) are used for descriptive, explanatory, and forecasting analyses of time series data in road safety research. Weijermars and Wesemann (2013) suggested that structural time series will further improve the forecasts – however, this is not technically feasible at present.

2.12.7 State-space model

In a recent study, Bijleveld et al. (2010) preferred state-space modelling based on a single multivariate model which consists of interpretable components. The missing observations are treated as part of the statistical analysis. However, it shall be noted that if the observations are obtained from credible traffic safety resources with high accuracy, then the use of univariate time series models may be sufficient to analyse the low counts.

According to Hermans et al. (2006), one of the most important characteristics of state-space models is that the observations are composed of distinct components such as trend, seasonal and regression elements, and each component is modelled separately with its own direct interpretation. These components are allowed to change with time. It shall be

noted that stationarity of the data series is not necessary if state-space models are used for time series analyses (Lu, 2007). Wesemann et al. (2010) used state-space modelling with stratified data on crashes (or victims) and mobility (in which the car crashes and car vehicle kilometres are extrapolated).

2.12.8 Model with stratified data

Broughton and Lawton (2009) prepared two series of forecasts and stratified the data according to road users and age groups in order to predict the rate of casualties in Northern Ireland. The data were stratified into five road user groups: (1) car occupants, (2) motorcyclists, (3) pedestrians, (4) pedal cyclists and (5) others. The data were also stratified into two age groups, comprising children (aged 0–15 years) and young adults (aged 16–24 years). Weijermars and Wesemann (2013) stratified road traffic fatalities according to three types of collisions: (1) car occupants, (2) fatalities among other road users that were hit by a car and (3) other fatalities. Stipdonk et al. (2010) used SWOV to stratify data according to traffic mode, age and type of road. Each method, however, results in a different expected number of casualties for year 2020.

Stipdonk et al. (2010) also highlighted that the three-step approach described by Broughton (2009) is a suitable method to forecast road safety. However, it shall be noted that there are several issues which need to be addressed before applying the three-step method. Firstly, one needs to determine how to extrapolate past data in order to forecast the rate of casualties for the targeted year. Secondly, it is implicitly assumed that the baseline trend of the rate of casualties contains the effects of numerous measures that are not explicitly included in the model. Thirdly, one needs to estimate the additional effects of specific road safety measures if a new policy is implemented. A generic approach has been proposed in order to address the issues above. The approach is based on stratification

(i.e. disaggregation) of crash data according to relevant factors such as traffic mode, age or the type of road. At present, there is no theoretical framework available that will facilitate the analyst in choosing the factors in which stratification will improve the model.

In order to decide on stratification for different crash groups, Stipdonk et al. (2010) considered two indications: (1) crash groups with different developments of crash rates with respect to time, and (2) crash groups with different crash rates, combined with different developments of mobility (shift in mobility between different crash groups with respect to time), as well as two conditions for stratification: (1) availability of stratified mobility data and (2) sufficient crash data.

In a recent work, Stipdonk et al. (2013) perceived that stratified risk analysis is relevant to analyse time series crash data. The overall risk (i.e. the number of total casualties divided by the total distance travelled) is influenced by travel modes and the type of roads. For instance, it is expected that the overall risk will be reduced by switching from unsafe travel modes to safer travel modes (public transport) and improving the conditions of roads. The overall risk is also influenced by changes in circumstances or safety measures which are typically applicable only to a subset of crashes. For these reasons, stratification is required to understand the effects of these changes on the overall risk. In the Netherlands, for example, it is not possible to stratify according to speed limits since such data are not recorded. Based on the few variables available in the Dutch travel survey, the data are stratified according to travel mode, age and gender, for the reasons mentioned previously.

In addition, Jacobs et al. (2000) suggested that the analyst should also consider underreported data when forecasting fatalities, especially in developing countries. According to Amoros et al. (2008), the reported number of fatalities in France is almost

complete – however, the reported number of non-fatal casualties is rather low, which is the case for most developed countries. Such data are heavily biased and therefore, there is a need for valid estimates. They found that the average annual estimated incidence is 3.4 times higher than that reported in police reports for all categories of injuries within the following study period: 1996–2004.

2.13 Key studies which involve the use of time series modelling to analyse and predict road safety

Road traffic accidents are recorded based on specific time intervals, which may be daily, weekly, monthly or annually. Thus, analysing road traffic fatalities requires analysing the changes in trends as a function of time. In addition, future road traffic fatalities can be related to past ones. For this reason, time series models have been used extensively to forecast the number and rate of fatalities.

A number of key studies which involve the use of time series modelling methods are summarized in Table 2.1.

Table 2.1: Key studies on road safety research which involve the use of time series modelling methods

| Researcher(s) | Objective/Context | Model used | Findings |
|------------------------|--|---|---|
| Ross (1973) | To evaluate the implementation of the Road Safety Act of 1967 which introduced scientific tests to determine and define the crime of drunk driving. | Interrupted time series analysis. | The Road Safety Act of 1967 did result in a fast and important decrease in traffic casualties, even though this effect may have been to a certain extent the result of the associated publicity campaign. |
| Wagenaar (1984) | To identify the underlying relationship between changes in the unemployment rate and involvement in motor vehicle crashes. | ARIMA and dynamic regression time series. | Vehicle miles travelled was not a significant intervening effect between unemployment and crash involvement. |
| Hilton (1984) | To evaluate the deterrent impact of road safety laws on the statewide incidence of fatal traffic accidents due to alcohol consumption. | Interrupted time series analysis. | The results do not show that a deterrent impact occurred among those accidents. |
| Campbell et al. (1986) | To analyse seat belt legislation in the US. | Intervention analyses. | Those results reveal about seven percent fewer fatalities among targeted occupants. |
| Scott (1986) | To analyse the relationship between the road accidents in Britain and a number of explanatory variables (traffic volume, petrol price, temperature, rain, workdays, oil crisis and speed limit). | Box-Jenkins. | The accident series are typically quite “noisy” and auto-correlation among the residuals from standard regressions not so strong, Box-Jenkins models are unlikely to represent the series appreciably superior to regression based on the assumption of uncorrelated residuals. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|----------------------------|---|---|--|
| Harvey and Durbin (1986) | To conduct an independent technical assessment on the effect of road safety laws on road casualties. | Intervention analyses. | There were substantial net reductions in numbers killed resulting from the introduction of the seat belt law. |
| Wagenaar and Maybee (1986) | To analyse the effects of raising the legal drinking age from 18 to 19 years in Texas. | Interrupted time series analysis. | It is clear that the 1-year rise in legal age in Texas had a substantial effect on youth crash involvement. |
| Wagenaar et al. (1988) | To examine the effects of compulsory seat belt use. | Box-Tiao intervention analysis time series. | A statistically significant decreased of 8.7% in the rate of front-seat fatalities in the first eight states with seat belt laws. |
| Ray (1989) | To evaluate the effectiveness of new drunken driving laws introduced in California. | Box-Jenkins. | A substantial decline in the fatality rates from car accidents following the enactment of the new drunken driving laws. |
| Wagenaar et al. (1990) | To examine the effects of the raised speed limits on injury morbidity and mortality. | Time series intervention analyses. | Significant increases in casualties on roads where the speed limit was increased. |
| Garber and Graham (1990) | To examine the effects of the new 65 mile-per-hour (mph) speed limit on the number of rural highway fatalities in the US. | Time series regression equations. | The new laws possess increased fatalities on both rural interstate and rural noninterstate highways in the majority of states, but also that these effects vary substantially across the states. |
| McKnight and Klein (1990) | To analyse accidents over the 1 year following the increase in the speed limit on rural Interstate highways. | Time series analysis. | Speeding on rural Interstates raised by 48% and fatal accidents raised by 27%. |
| Asch et al. (1991) | To investigate the effectiveness of New Jersey's mandatory belt use law by testing for risk-compensation effect. | Box-Jenkins intervention analysis. | The actual safety effect of the law was probably diluted by risk-compensating behavior. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|----------------------------|---|------------------------------------|--|
| Fleming and Becker (1992) | To study the effect of implementing mandatory motorcycle helmet law for all operators and passengers. | Box-Tiao time series intervention. | The decreases of 12.6% and 57.0% were projected for total and head-related fatalities during the year after the law was implemented. |
| Murry et al. (1993) | To examine the impact of paid advertising campaigns on the drunk driving behaviour of male youth. | Time series intervention. | The advertising campaign lowered youthful male drinking and driving behavior and, accordingly, traffic accidents. |
| Neustrom and Norton (1993) | To analyse the effect of stricter drunk driving laws. | Interrupted time series analysis. | The new law deterred drinking and driving. |
| Ross and Klette (1995) | To examine the adopted legal reforms, including the abandonment of mandatory jail sentences for individuals driving with BACs above the specified limits. | Interrupted time series analysis. | The traffic deaths declined simultaneously with the reforms, conforming to the understanding that Scandinavian success in lowering impaired driving does not rely on mandatory jail. |
| Ballart and Riba (1995) | To evaluate the effects of deterrence-based reforms on the use of helmets by motorcyclists and moped riders. | Interrupted time series analysis. | Substantial decline in the number of serious cases in the treated group smaller motorcycles and mopeds - a reduction which does not occur in the control group - larger motorcycles. |
| Ledolter and Chan (1996) | To examine whether there is a significant change in the rate of fatal and major-injury accidents following the implementation of higher speed limits. | Time series intervention. | A 20% rise in the number of state-wide fatal accidents to the speed limit change. The effect is largest on rural interstates. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|----------------------------|---|--|--|
| Williams et al. (1996) | To examine the effect of implementing a programme whose objective is to increase seat belt and child restraint use and to decrease other traffic law violations, including alcohol. | Structural time series. | It is expected that 45 more fatalities and 320 more serious injuries would have occurred during the 6 months following the program than were actually identified. |
| Jessie and Yuan (1998) | To examine the determinants of road traffic fatalities in Singapore, with emphasis on seat belt regulations, use of breathalysers, as well as circuit training and testing systems. | Time series regression. | The seat belt legislation does not have any effect on fatalities. The use of breathalysers was seen to be effective in decreasing occupant fatalities. The circuit training and testing system was found to be effective in decreasing fatalities. |
| Farmer et al. (1999) | To predict the number and rate of fatalities. | Time series cross section regression. | Fatalities on interstates higher 15% in the 24 states that raised speed limits. |
| Beenstock and Gafni (2000) | To propose a theoretical model whereby the road safety is dependent e.g. police enforcement and technology. | Smeed's model and vector autoregression (VAR). | The downward trend in the road accident rate is part of a global phenomenon. |
| Noland (2001) | To examine the effect of various highway improvements on both fatalities and injuries. | Cross-sectional time series. | Strongly refute the hypothesis that engineering design improvements are beneficial for decreasing total fatalities and injuries. Whereas raises in medical technology have accounted for a significant share of overall reductions in fatalities. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|-------------------------------|---|-----------------------------------|---|
| Oppe (2001) | To investigate and forecast the developments of traffic and traffic safety. | ARIMA and structural time series. | Forecasts based upon the model assumptions reveal for the near future a more rapid decrease in risk than before. |
| Bédard et al. (2001) | To forecast the number of older drivers and passengers who may be fatally injured in traffic crashes. | Least squares regression models. | The number of fatally injured women and men aged 65 and older will rise respectively by 373% and 271% between 1975 and 2015. |
| Lassarre (2001) | To compare the road safety developments in 10 European countries. | Structural model. | Europe's road systems can thus absorb a 6% rise in traffic annually while maintaining the number of fatalities constant. |
| Houston and Richardson (2002) | To explore whether changes in the existing seat belt laws from secondary to primary enforcement will improve safety. | Box–Tiao intervention analysis. | California encountered an improvement in traffic safety in terms of a substantial decrease in injuries, but the change in enforcement provision had no statistically significant effect on fatalities. |
| Gaudry (2002) | To explain what is road use demand, accident and the severity of accidents, with focus on the number of victims, and to explore the influence of various factors. | DRAG. | Explanatory factors considered: demand, prices, vehicles (quantities and characteristics), legal regimes and police, service levels of modes, infrastructure quality), general, age, sex, vigilance, economic activities and trip purposes. |
| Hess and Polak (2003) | To analyse the effects of Speed Limit Enforcement Cameras (SLEC) on the rate of accidents in Cambridgeshire. | ARIMA Model. | The result showed a reduction in injury accidents by 31.26%, thus providing clear evidence that SLECs do in fact lead to a significant decline in accident numbers. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|-----------------------|--|--|---|
| Kopits (2003) | To investigate the basic relationship between road safety and income levels. | Quadratic and spline models with region-specific log-linear time trends. | Prediction of future traffic fatalities shows that the global road death toll will increase by approximately 66 percent between 2000 and 2020. |
| Bossche et al. (2004) | To develop models which will forecast the frequency and severity of accidents in Belgium. | ARIMA | A substantial impact of weather conditions and laws and regulations on traffic safety, but there appears to be negligible statistical impact of economic conditions. |
| Vernon et al. (2004) | To analyse the effects of increased speed limits on the rate of crashes, fatality crashes and injury crashes on highways in Utah. | Intervention time series analysis. | An adverse impact on crash occurrence for subsets of crash types and highways, but do not show a major overall effect of speed limit law repeal and raised speed limit on crash occurrence. |
| Muller (2004) | To examine the effect of exempting adult motorcyclists and moped riders from wearing helmets in Florida. | Interrupted time series analysis. | Recommend that the law's edge exemption should be revoked. |
| Noland (2004) | To analyse the impact of changes in the average fuel efficiency on traffic-related fatalities while controlling for other confounding effects. | Negative binomial regression. | There may have been a relationship between fleet fuel efficiency improvements and traffic fatalities in the 1970s, this has largely disappeared. |
| Hermans et al. (2005) | To study the impact of meteorological, socioeconomic, legislative and calendar factors on the exposure and risk. | Unobserved components methodology. | Precipitation and snow increase accident risk while temperature plays a substantial role for exposure. The economic indicators mainly affected accident risk. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|-----------------------------|---|-------------------------------------|--|
| Law et al. (2005) | To examine the effects of both motorcycle safety intervention and economic crisis on motorcycle-related fatalities in Malaysia. | Transfer function. | Motorcycle Safety Programme has been one of the effective measures in decreasing motorcycle safety problems in Malaysia. The performance of the country's economy was also found to be substantial. |
| Kim et al. (2006) | To evaluate the effects of five traffic safety policies implemented in Korea. | ARIMA intervention impact analysis. | No statistically significant abrupt and permanent impacts of the five policies on decreasing the numbers of total crashes, fatalities and injuries were found. |
| Hermans et al. (2006) | To investigate the monthly frequency and severity of road traffic accidents in Belgium. | State-space methods. | Laws relating to seat belts, speed and alcohol have proven successful. Frost enhances road safety while sun has the opposite effect. Precipitation and thunderstorm specifically influence accidents with light injuries. Economic conditions have a limited impact. |
| Krystek and Żukowska (2007) | To examine the correlation between the number of traffic fatalities and the degree of public activity. | Structural time series. | The falling public activity decreased the demand for transport leading to a decline in miles traveled. This in turn resulted in fewer accident fatalities. |
| Bossche et al. (2007) | To analyse the yearly number of fatalities per age and gender group of Belgians. | State-space. | Road risk is changing over the age groups according to a U-shaped curve, and that men typically have a higher risk than women. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|--------------------------------|--|---|---|
| Males (2007) | To analyse the effects of California's strict Graduated Driver Licensing (GDL) Law on the deaths of drivers aged 16–19. | Incidence rate ratios and ARIMA. | California's GDL may negatively affect older teenagers and other driver subpopulations and merits reevaluation. |
| Wagenaar et al. (2007) | To examine the effects of changing the legal BAC limits in 28 states in the US. | Interrupted time series quasi-experimental. | Changes in legal BAC limits substantially affected alcohol-related fatal crash involvement. |
| Bijleveld et al. (2008) | To introduce a general multivariate model for 'event risk' analysis that accounts for exposure, risk and severity. | Latent risk time series. | It is found that the general methodology can be effectively applied in the assessment of risk. |
| Richardson and Shaw (2009) | To analyse the impact of vehicle concentration, unemployment, length of roadways, legal driving blood alcohol levels, speed on roadways, gross domestic product per capita, and alcohol consumption. | Cross-sectional time series regression. | The key finding is that maximum speed limits on highways significantly affect traffic safety, and reducing these limits would decrease traffic fatalities. More alcohol consumption is significantly related to a higher traffic fatality rate. |
| Bijleveld et al. (2010) | To develop a yearly time series on the number of fatal accidents (inside and outside urban areas) and the number of kilometres driven by motor vehicles in the Netherlands. | Multivariate time series model. | Prefer the state-space analysis which is based on a single multivariate model that consists of interpretable factors which can treat missing observations as part of the statistical analysis. |
| Al Bavon and Standerfer (2010) | To determine the effect of the Texas motorcycle helmet law on the number of fatalities. | ARIMA. | The repeal of the universal helmet law in Texas in 1997 has had a substantial adverse effect on motorcyclist fatalities. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|---------------------------|---|---|---|
| Gargett et al. (2011) | To examine the progress made in the states and territories of Australia in terms of the reduction in fatalities. | Univariate time-series. | Statistically significant declines in fatality rates since 1971 were found for all jurisdictions with the national rate reducing on average, 3% per year since 1992. |
| Yannis et al. (2011a) | To provide a parsimonious model which links the motorization level with the decrease in rate of fatalities observed across EU countries. | Autoregressive non-linear. | An autoregressive log-transformed model appears to outperform the base autoregressive non-linear model in this respect. |
| Izquierdo et al. (2011) | To analyse the contribution of the penalty point system (which is the most important legislative measure for driving licences) in reducing road traffic fatalities. | ARIMA and intervention model. | The combination of three elements: the penalty point system, the gradual stepping up of surveillance measures and sanctions, and the publicity given to road safety issues in the mass media seem to be the key to success. |
| Bergel-Hayat (2012) | To describe how time series analysis of road risk is carried out at the national level in Europe. | Various time series models. | Even though descriptive and explanatory models differ, both types can be used for providing forecasts. |
| Commandeur et al. (2013) | To demonstrate the applicability of standard methods of statistical inference on the impact of road safety violations. | ARMA, DRAG, structural time series and state-space methods. | The methods result in underestimation or overestimation of standard errors, which lead to erroneous inferences. |
| Gaudry and Himouri (2013) | To construct the first countrywide model in Algeria. | DRAG | The demand for road use and of road safety outcomes for Algeria. |

Table 2.1, continued

| Researcher(s) | Objective/Context | Model used | Findings |
|----------------------------|---|--------------|---|
| Antoniou and Yannis (2013) | To analyse macroscopic road safety trends in Greece. | State-space. | Demonstrates the benefits of employing advanced state-space modeling techniques for modeling macroscopic road safety trends. |
| Izquierdo et al. (2013) | To analyse the influence of factors such as socioeconomic, environmental, technological and legislative factors of accidents with injured victims and fatalities occurring on interurban roads. | DRAG. | The greatest influence revealed by the results is exposure as well as inexperienced drivers, speed and an ageing vehicle stock, have a negative effect, while the increased surveillance on roads, the improvement in the technological features of vehicles and the proportion of high capacity networks have a positive effect. |

2.13 Recent time series models used in road safety

One of the modelling approaches that have been frequently used to model a control process as an intervention is the interrupted time series technique. This technique is used to evaluate changes in the accident levels. More sophisticated versions of interrupted time series technique involve the use of univariate ARIMA models (Votey, 1986). In general, time series models are better than cross-sectional models because cross-sectional models do not account for geographical and cultural differences between countries, states or provinces (Page, 2001). Econometric models such as regression and time series models are very useful to enhance one's understanding on traffic safety trends (Bossche et al., 2004). The progress of time series analysis techniques as well as the availability of more detailed time series data has led to advances in statistical approaches in road traffic safety (Bijleveld et al., 2006).

Moreover, an important class of road safety models is based on time series analysis. A time series basically consists of repeated measurements of a certain phenomenon at regular intervals, and therefore, the evolution of road safety with respect to time can be studied in this manner (Hermans et al., 2005). Time series can be defined as a stochastic process in which the subsequent observations randomly change with time. The observations may be the number of people killed or the number of people injured in road traffic accidents, or the total number of accidents over a month, quarter or year. These data are used to create the time series which are then used to build the model (Krystek & Żukowska, 2007). ARIMA intervention analysis has been used in numerous studies pertaining to the assessment of traffic safety policies (Kim et al., 2006). However, Box-Jenkins methods are more widely used for time series analyses compared to state-space methods (Hermans et al., 2006).

According to Hermans et al. (2006), one of the most important characteristics of state-space models is that the observations are composed of distinct components such as trend, seasonal and regression elements, and each component is modelled separately with its own direct interpretation. These components are allowed to change with time. It shall be noted that stationarity of the data series is not necessary if state-space models are used for time series analyses (Lu, 2007). Both ARIMA and DRAG are dedicated time series models which account for dependencies between the observations in the time series data (Reurings & Commandeur, 2007). However, Bijleveld et al. (2008) highlighted that structural time series models are preferable from a methodological perspective.

In a recent study, Bijleveld et al. (2010) preferred state-space modelling based on a single multivariate model which consists of interpretable components. The missing observations are treated as part of the statistical analysis. However, it shall be noted that if the observations are obtained from credible traffic safety resources with high accuracy,

then the use of univariate time series models may be sufficient to analyse the low counts. According to Broughton and Knowles (2010), various approaches have been developed to evaluate the developments in road traffic safety in different countries, and some of these methods have been used for forecasting. The ARIMA and DRAG2 models have been used in Belgium, Canada, France and Sweden. However, structural time series models are the preferred models in the Netherlands.

Yannis et al. (2011a) developed parsimonious models which link the motorization level with the decrease in rate of fatalities across countries in the EU in the past three decades. They observed that the autoregressive log-transformed model outperforms the base autoregressive non-linear model, and enables them to identify the countries with the best and worst road safety performance. Yannis et al. (2011a) also highlighted that time series methods have been used to remove temporal trends and autocorrelations from traffic fatality risk models, which enables one to assess the impact of macroscopic road safety-related model parameters on traffic risk. According to Sukhai et al. (2011), NARX models are similar to ARIMA models that have been used extensively in road safety research. Such models do not only contain autoregressive terms of the variable being studied, but also exogenous input terms that can be either autoregressive or obtained from the present time step.

According to Bergel-Hayat (2012), autoregressive models as well as those that are similar to autoregressive models once the observed data have become stable are the models commonly used for structural modelling. The most advanced explanatory autoregressive model for road risks was developed by Gaudry in 1984.

According to Antoniou and Yannis (2013), it is recommended that time series methods such as ARMA-type models (stationary data) and ARIMA or state-space models (non-stationary data) are used for descriptive, explanatory, and forecasting analyses of time

series data in road safety research. Weijermars and Wesemann (2013) suggested that structural time series will further improve the forecasts – however, this is not technically feasible at present.

2.14 Statistical methods used to forecast time series data

In general, forecasting methods are classified into qualitative and quantitative methods. In qualitative methods, the forecasts are often made based on the analyst's judgement, which relies heavily on the analyst's intuition and experience. In contrast, quantitative methods are based on mathematical or statistical models. In this study, quantitative methods are used since they are commonly used in road safety research.

Commandeur et al. (2013) reviewed a number of time series analyses on national road safety trends in Europe since the 1980s. Time series models have progressed from descriptive towards explanatory models, as well as from deterministic towards stochastic models under the form of structural models. In general, traditional regression models (linear, generalized linear and non-linear models) are unable to capture the dependencies inherent in time series data.

2.14.1 Univariate and multivariate methods

There are two types of quantitative forecasting methods: (1) multivariate methods which includes explanatory factors and (2) univariate methods which exclude explanatory factors and the forecasts are made based on the values of the response variables. In multivariate methods, the forecasts of a variable depend, at least partly, on the values of one or more additional time series variables (predictor or explanatory variables). If the variables are dependent one way or another, the multivariate forecasts may require the use of a multivariate model that is built based on a number of equations. In contrast,

univariate forecasting methods are merely dependent on the present and past values of the single series being forecasted, and may be augmented by a time function such as a linear trend (Chatfield, 2000).

2.14.2 Deterministic and stochastic models

Quantitative models can also be classified as deterministic or stochastic (probabilistic) models. The relationship between the dependent variable (Y) and the explanatory or predictor variables (X) of a deterministic model and stochastic model is shown in Equation (2.1) and Equation (2.2), respectively. Deterministic models are usually used in the physical sciences since the function f and coefficients β_1, \dots, β_m are known for certain. The relationship of Y and X in the social sciences, however, is usually random (stochastic). It shall be noted that measurement errors as well as the variability of other uncontrolled variables will introduce stochastic components into the relationship. The error components are a realization from a certain probability distribution and are generally termed as ‘noise’. In most cases, the functional form f and the coefficients are unknown and therefore, they need to be determined from past data. The data are typically presented in the form of time-ordered sequences, which are known as time series. The statistical models in which the available observations are used to determine the model representation are also known as empirical models (Abraham & Ledolter, 2005).

$$Y = f(X_1, \dots, X_p; \beta_1, \dots, \beta_m) \quad (2.1)$$

$$Y = f(X_1, \dots, X_p; \beta_1, \dots, \beta_m) + \text{noise} \quad (2.2)$$

2.14.3 Microscopic and macroscopic levels

The scope of the forecasts is also critical in road safety research. Most of the studies in this area fall under the microscopic level, and it can be observed from the literature that the implementation of road safety policies, geometric characteristics and road elements has been analysed extensively over the years. The forecasted number of road traffic casualties at the macroscopic level is usually less accurate and therefore, forecasts made at the microscopic level are preferable over those at the macroscopic level. There are various factors which influence the road safety characteristics in a jurisdiction – however, the lack of reliable data as well as inconsistencies in past research findings indicate that there are still uncertainties in this area which need to be resolved.

2.14.4 Time series data patterns

There are four general types of time series data patterns: (1) horizontal, (2) trend, (3) seasonal, and (4) cyclical. A horizontal pattern can be observed when the data observations fluctuate around a constant level or mean. This pattern indicates that the data series is stationary in its mean. A *trend* pattern is apparent when the data observations increase or decrease over a time period. A *cyclical* pattern is evident when the observations increase and decrease erratically within the time period. The cyclical component is indicated by wavelike fluctuations around the trend. A *seasonal* pattern exists when the observations are influenced by seasonal factors. The seasonal component refers to a change in pattern that repeats itself year after year (Hanke & Wichern, 2005).

2.14.5 Models used for analysing and forecasting road traffic casualties

The time series models used to forecast road traffic casualties as presented in the COST 329 (2004) report are listed as follows:

- (1) Dynamic univariate models such as deterministic component and ad hoc models, ARIMA models, structural models, state-space models and autoregressive Poisson models i.e., based on log-linear transition models.
- (2) Explanatory models such as linear and non-linear models, transfer function-noise models and intervention analysis.
- (3) VAR and simultaneous equation models.

The report provides the following conclusions and recommendations regarding the use of time series models in forecasting road traffic casualties:

- (1) Models that are based on yearly or monthly number of accidents or casualties such as polynomial spline models are not recommended for use as forecasting tools. ARIMA models may be used for forecasting purposes up to two years.
- (2) Models that are based on yearly or monthly vehicle kilometres as a measure of exposure are suitable for forecasting up to 10 years. Deterministic risk models fall in this category, and it is preferable that the analyst uses the exponential risk model with or without interventions. Stochastic risk models, structural models (Harvey), as well as extended exponential risk models are also suitable for forecasting up to 10 years.
- (3) Models which make use of additional variables such as simple regression models (mostly of log-linear type) with up to four extra variables can be used for forecasting up to five years. ARIMAX models with calendar effects and some interventions can also be used for this purpose.

- (4) Model which make use of medium to large numbers of explanatory variables for forecasting such as DRAG-type econometric models (e.g. TAG or TRULS) are suitable for long-term forecasting. Multivariate ARMAX or structural (Harvey) models are also used for this purpose. However, it shall be noted that multivariate ARMAX models are only suitable for short-term and medium-term forecasts up to two years.

2.15 Time series models used in this study

Considering the suggestions as well as conclusions provided by the researchers as previously discussed, has concluded that three time series models that will be used in this study are: (1) Box-Jenkins (ARIMA) model, (2) transfer function-noise model, and (3) state-space model of the structural models. Below are the researchers that are strongly supporting the three models as mentioned above.

- (1) Hauer (2010) stated that it is not clear as to whether the predictions made based on data in statistical regression (structural) models are superior to those based on extrapolating time trends (univariate models).
- (2) ARIMA intervention analysis has been used in numerous studies pertaining to the assessment of traffic safety policies (Kim et al., 2006).
- (3) ARIMA methods are more widely used for time series analyses compared to state-space methods (Hermans et al., 2006).
- (4) Bijleveld et al. (2010) preferred state-space modelling based on a single multivariate model which consists of interpretable components. If the observations are obtained from credible traffic safety resources with high accuracy, then the use of univariate time series models may be sufficient to analyse the low counts.

- (5) Commandeur et al. (2013) described that time series models have progressed from descriptive towards explanatory models, as well as from deterministic towards stochastic models.
- (6) Antoniou and Yannis (2013) recommended that time series methods such as ARMA-type models (stationary data) and ARIMA or state-space models (non-stationary data) are used for descriptive, explanatory, and forecasting analyses of time series data in road safety research.

In summary, the three models used can be classified into:

- Box-Jenkins (ARIMA) model is under quantitative, univariate and stochastic method.
- Transfer function-noise model is under quantitative, univariate and stochastic method.
- State-space model is under quantitative, multivariate and deterministic method.

2.15.1 Box-Jenkins (ARIMA) models

A basic requirement for ARIMA models is that the time series must be stationary. ARIMA models are developed based on the assumption that there is a stable relationship between the observation at time t and previous observations. Hence, the values at succeeding points in time are correlated. The original series is then duplicated and shifted one place upwards. If the correlation coefficient is significantly different from zero, then serial correlation exists. This process is repeated for a further shift. The ARIMA model consists of two main parts. The first part is the autoregressive (AR) model, which directly models the autoregression. The second part is the moving average (MA) model, which is basically a complementary model. It is not defined based on a direct relationship between succeeding observations – rather, it is defined based on their errors. However, it shall be

noted that each model can be translated into other approaches. Selection of the most efficient approach is generally dependent on the practical characteristics of the series.

The ARIMA model is basically an integration of autoregressive and moving average models, and it is commonly used because of its flexibility. The letter 'I' which lies in the middle of the name 'ARIMA' stands for integration if a difference operator is needed in order to make the series stationary (COST 329, 2004). The ARIMA model is a very powerful tool which yields accurate short-range forecasts in time series analyses. In general, ARIMA models are quite flexible and can be used to represent a wide range of characteristics of time series that occur in practice. Formal procedures are available in order to test the adequacy of the models. The forecasts and prediction intervals follow directly from the fitted model. The methodology used for ARIMA modelling will be described in detail in the following chapter.

2.15.2 Intervention analysis and transfer function-noise models

Intervention analysis is a formal test used to determine variations in the mean of a time series due to an indicator variable which contains discrete values that flag the occurrence of an event affecting the response series. In the absence of an intervention, the series is usually assumed to be a pure ARIMA process or structural model. ARIMA models represent a general class of models that can be used very effectively in time series modelling and forecasting problems. It is implicitly assumed in these models that the conditions remain the same for data of a time series process. However, if the conditions vary with respect to time, the ARIMA models can be improved by introducing certain inputs which reflect the changes in the process conditions. This leads to the formulation of *transfer function-noise models*. These models can be viewed as regression models with

a serially dependent response, as well as inputs and error terms. Identification and estimation of these models can be challenging (Montgomery et al., 2008).

In addition, forecasting future observations relies on past data. It is inherently assumed that the conditions during which the data are collected will remain the same in the future. In many situations, however, this assumption may be invalid. In such cases, the use of intervention models such as transfer function-noise models is appropriate. A set of input variables which will have an effect on the time series are added into the model, which offers more suitable options.

2.15.3 State-space models

State-space models are based on the Markov property, which implies the independence of the future of a process from its past, given the state of the present system. In this system, the state of a process at the current time contains all of the past information that is required to predict the process behaviour in the future (Montgomery et al., 2008). Yaffee and McGee (2000) defined this model as jointly stationary multivariate time series processes that have dynamic interactions, and the model is constructed from two basic equations. The state transition equation consists of a state vector of auxiliary variables which is a function of a transition matrix and an input matrix, whereas the measurement equation consists of a state vector that is canonically extracted from observable variables. These vector models are estimated using a recursive protocol and can be used for multivariate forecasting. According to Brockwell and Davis (2002), these models make up an extremely rich class of time series models which are superior to linear ARIMA and classical decomposition models.

2.15.4 Least squares and maximum likelihood methods

Model selection and fitting involves choosing one or more forecasting models and fitting the model to the data. Fitting or estimating the unknown model parameters is usually done using the *ordinary least squares* (OLS) method. The least squares method selects the model parameters by minimizing the sum of squares of the errors (ε_i). This estimation method is commonly used in analysis of variance and regression analysis. *Maximum likelihood* (ML) has also been widely used. According to Yaffee and McGee (2000), this approach involves asymptotic iterative estimation of parameters by maximizing the likelihood that the model will fit the data. This procedure involves modelling the likelihood, taking its natural log, finding its minimum, and estimating the parameter values that minimize the lack of fit.

2.16 Chapter summary

In the recent years, various variables related to fatalities involves: population, motorized vehicle fleet per capita, percentage of urban population, percentage of population aged 15-24 years, percentage of population who are active and employed, consumption of pure alcohol per capita and percentage of buses and coaches in the motorized vehicle fleet. Generally, the variables related to road safety could be classified into: (1) socioeconomic and demographic indicators, (2) age, gender categories and urbanization, (3) weather and time variations, (4) medical care facilities, (5) geographical factors and road geometry, (6) proportion of motorcycle and public transit user and (7) other explanatory variables such as the human development index.

In the literature, the road safety measures implemented involves: (1) seatbelt and other motor vehicle safety standards, (2) maximum blood alcohol concentration, (3) speed

limit, (4) speed limit enforcement camera, (5) provision for motorcyclist, (6) driver licensing and (7) laws enforcement.

There are various unit of exposure to risk used in road safety. Wegman and Oppe (2010) suggested that the use of mortality rates (the number of fatalities per head of the population) should be used to compare traffic risks – however, the use of mortality rates is negated by the fact that the motorization level is not taken into account. For this reason, the fatality risk is commonly used as a criterion to assess road traffic safety, and it is defined as the number of fatalities per motor vehicle kilometre. However, it shall be noted that if the data on motor vehicle kilometres are unavailable, then the rate of fatalities should be used, which is defined as the number of fatalities per motor vehicle.

In general, the model used in road safety research involves: (1) exponential model, (2) extrapolation model, (3) structural time series model, (4) time series analysis and intervention model and (5) state-space model. Considering the suggestions as well as conclusions provided by researchers, three time series models will be used in this study are: (1) Box-Jenkins (ARIMA) model, (2) transfer function-noise model, and (3) state-space model of the structural models. The classification of the three models used are: (1) ARIMA model is under quantitative, univariate and stochastic method, (2) transfer function-noise model is under quantitative, univariate and stochastic method and (3) state-space model is under quantitative, multivariate and deterministic method.

CHAPTER 3: RESEARCH METHODOLOGY

The methodology adopted in this study is described in detail in this chapter. This chapter begins with a description on the data and tools used in this study, followed by a description on the forecasting methods. The methodology used to develop each of the forecasting models is elaborated in detail, which includes the mathematical representations, parameters and criteria used to select the best model for forecasting. Multiple regression, ARIMA, transfer function-noise and state-space models are considered in this study.

The multiple regression is used to select the explanatory variables that are correlated significantly with the number of road traffic fatalities. Then, these variables are the input variables for the state-space model. The effectiveness of a road safety measure is checked with the ARIMA and transfer function-noise model. Whilst, for forecasting the fatalities up to year 2020, the ARIMA, transfer function-noise and state-space models are employed, respectively.

3.1 Construction of models

One of the goals in this study is to develop time series models based on the available models discussed in Chapter 2 in order to analyse and predict the rate of road traffic fatalities in Malaysia. The road safety characteristics are also explored in order to determine the impact of various road safety measures and compare the results with the findings of previous studies. In order to achieve this goal, a number of statistical techniques are used which will be described in detail in this chapter. A comprehensive review of relevant literature is carried out in the preliminary stages of this study, followed by data collection. Statistical analysis is used to determine the descriptive statistics of

road traffic casualties, which serves as the foundation for data stratification according to the type of road users and road environments.

The next step involves modelling the series of road traffic fatalities using univariate and multivariate analysis methods in order to determine the statistically significant variables which influence the rate of road traffic fatalities in Malaysia and determine the impact of the interventions (road safety measures) along with the forecasted rate of fatalities.

Based on the review of relevant literature in Chapter 2, there are 15 explanatory variables (independent variables) that are expected to influence the rate of fatalities, i.e. GDP per capita, hospital beds per 1,000 people, population, male population, female population, population age 15–24 years old, registered vehicles, registered motorcycles, registered buses, registered motorcycles, road sector energy consumption, road length, rural population, urban population and unemployment rate. Thus, all these 15 variables will be investigated further in this study using stepwise regression to obtain its significance.

The modelling process consists of the following steps: (1) model construction and evaluation, (2) model extrapolation and (3) forecast evaluation. The methodological framework of the modelling process in this study is shown in Figure 3.1 as follows.

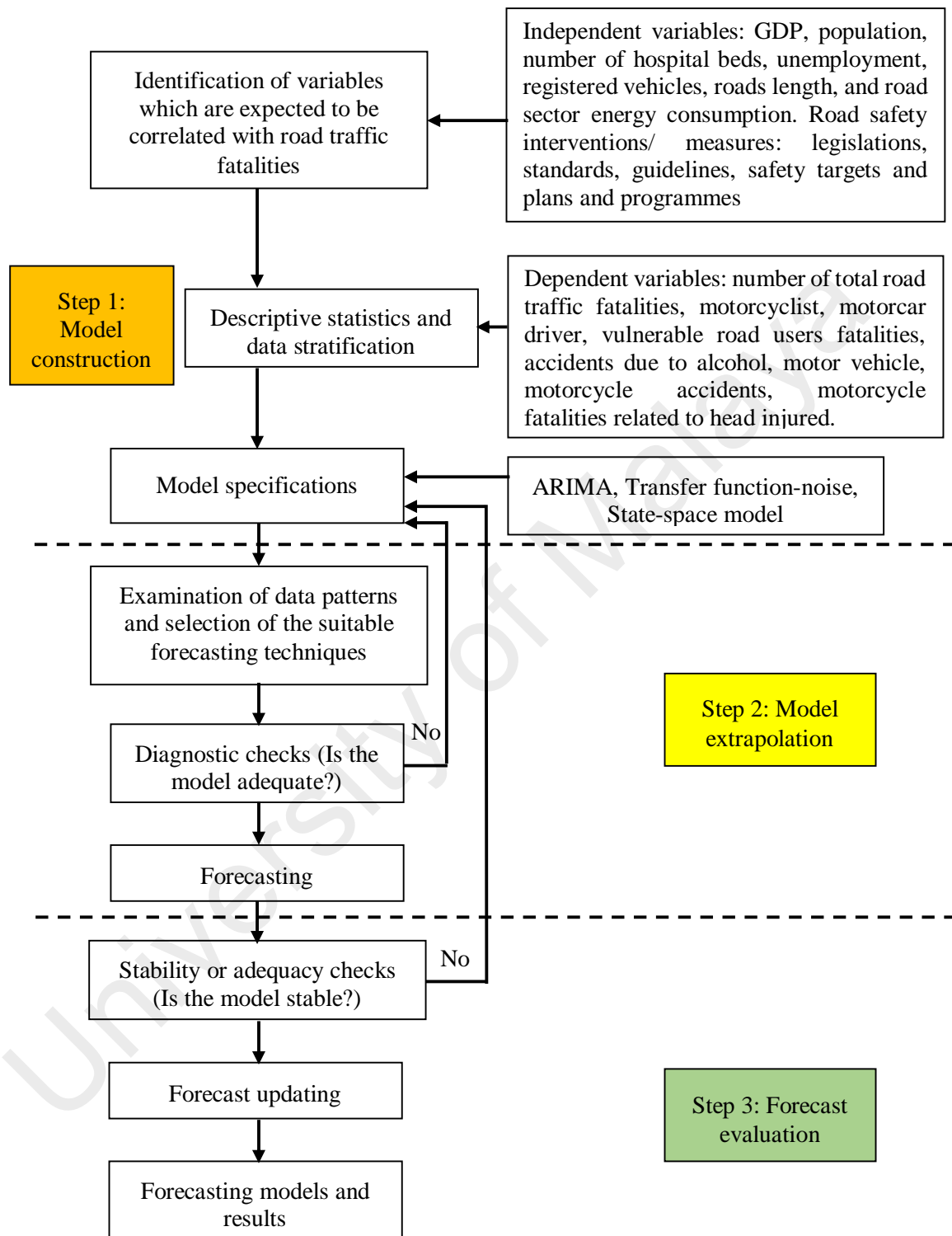


Figure 3.1: Framework of selected time series modelling

3.2 Research data and its reliability

Various types of data related to the exposure, risk and loss of road traffic accidents are employed in this study (Table 3.1). These data comprise the historical number of road traffic casualties according to the types of road users, demographic data, number of registered vehicles, socioeconomic status, infrastructure and weather data. The data used in this study are obtained from five sources which include: (1) the Royal Malaysian Police (PDRM), (2) Department of Statistics under the Prime Minister's Office of Malaysia, (3) Ministry of Transport, (4) Ministry of Works and (5) the World Bank. In terms of the accuracy, the data are good as the data are published officially by the recognized agencies. In fact, WHO estimations on the number of road traffic fatalities are similar with the number reported by the police.

Table 3.1: Data and its source

| Data | Sources |
|--|--------------------------|
| Number of total road traffic fatalities, motorcyclist fatalities, motorcar driver fatalities, vulnerable road users fatalities, accidents due to alcohol, motor vehicle accidents, motorcycle accidents, motorcycle fatalities related to head injured | Royal Malaysian Police |
| Total population, male, female, urban, rural population, population age 15–24 years old, number of hospital beds, unemployment | Department of Statistics |
| Number of registered vehicles (total), registered motorcycles, registered buses | Ministry of Transport |
| Roads length | Ministry of Works |
| GDP, road sector energy consumption | the World Bank |

There are a couple of constraints since this study deals primarily with historical data. The first constraint is that the data published during the early years of the period of investigation are not similar to those published in recent years. The second constraint is

that the available data are often published annually and it is generally more desirable to use monthly data in time series analysis. Even though monthly data consist of a higher number of observations, these data will elicit information regarding the variations in the data series. The explanatory variables that are expected to influence the number and rate of road traffic casualties are also published annually. However, a number of researchers have conducted road safety research based on annual data series. This is the typically the case for low-income and medium-income countries.

In Malaysia, road traffic safety data are collected, recorded and published by PDRM. A standardized form named 'Pol. 27' is used by the police to collect data related to road traffic accidents at accident scenes. The data consist of the following information: time and type of accident, the number of vehicles involved, road environment, injuries of vehicle occupants (if any) as well as the location of the accident. The reports published include the data and statistics on road traffic safety. However, the data published from the late 1960s to 1980 are only provided for Peninsular (West) Malaysia and the data for East Malaysia (Sabah and Sarawak) are provided beginning from 1981 onwards.

The road safety data for West and East Malaysia are available separately until 1988 and have been integrated since then. In the early years, the reports only provide the number of vehicles involved in accidents, the main causes of the accidents, the number of casualties, the accidents and casualties according to the type of road environments, demographic data of road users, the type of vehicles commonly involved in accidents, analysis of the investigations, and revenue from traffic cases. Generic data such as population, the number of registered vehicles and length of roads are also included. Statistics on road safety such as the rate of accidents and the rate of casualties are generated based on these data. However, the reports in recent years include details on motorcyclist and pedestrian casualties.

Attempts have been made by road safety authorities to improve the statistics on road safety in the reports. One such example is the reports published in 1978, in which the number of accidents and casualties for motorcyclists are more detailed. The rate of accidents and rate of casualties are also provided in such reports since year 1981. In 1988, charts are made available in the reports. In 1992, more comprehensive data and statistics on motorcyclist and pedestrian accidents are included. In 1998, accident data are divided into two main parts: Part A and Part B. Part A contains statistics on general road accidents derived from the monthly returns sent by all district police stations (traffic branches). Part B contains records on road accidents and statistics generated by a computer programme known as Microcomputer Accident Analysis Package (MAAP). The data used in this software are obtained from the revised standard road accident information form, i.e. Pol. 27 Pin. 1/91. The definitions of the terms used in the reports are provided along with notes to users.

The demographic, socioeconomic and weather historical data are extracted from the annual reports published by the Department of Statistics, Prime Minister's Office of Malaysia as well as from the online database of World Bank (2014). The number of registered vehicles are obtained from the reports published by the Road Transport Department, Ministry of Transport whereas the length of the roads are obtained from reports published by the Highway Planning Unit under the Ministry of Works. In this study, it is generally preferable to use future data estimated by national and international official agencies – however, if these data are not available, such data will be predicted using the models developed in this study.

Furthermore, during the data analysis in this study, the detailed of road traffic casualties data are available up to year 2012. For this reason, the descriptive statistics pertaining the road safety are conducted using year 2012 data. To develop time series

models to predict the fatalities, number of fatalities from 1981 to 2012 are used to estimate the models. Whilst, to evaluate the forecasts, number of fatalities in 2013 and 2014 are used.

3.3 Data analysis tools

A number of software packages are used to analyse the data in this study, namely Microsoft Excel (2010), Minitab (Version 16), Statgraphics and SAS (Version 9). Microsoft Excel is a spreadsheet application supplied with Microsoft Office Suite, which provides tools to present and organize data, compute statistical data and plot graphs, and it also offers macro programming language. Minitab, Statgraphics and SAS are leading software packages in statistical analysis. Minitab is convenient for univariate analysis whereas SAS package is capable of performing numerous statistical operations in this study. In addition, EvIEWS (Version 6) is used to crosscheck the unit root test results and is more likely to produce consistent results.

3.4 Selection of a forecasting method

Since one of the objectives of this study is to analyse the impact of road safety measures that have been implemented in Malaysia as well as to forecast the rate of road traffic fatalities, it is crucial to select the most appropriate forecasting methods that will fulfil this aim. A number of factors have been considered, taking into account the findings of previous studies. For this reason, quantitative methods are used based on forecasts at the macro level. The explanatory variables and interventions (road safety measures that have been implemented in Malaysia) are incorporated into the models. It is important that the explanatory variables are incorporated into the models to ensure that the predictions generated by the models are not an extension of the current trend of casualties in

Malaysia. Malaysia has yet to experience a single peak followed by a decrease in the absolute number of fatalities. The use of a univariate model will result in a prediction in the rate of road traffic fatalities that is constantly increasing in the future, which is contradictory to the trend in developed countries. The forecasting period is chosen to be 8 years (up to year 2020), which can be classified as ‘medium-term forecast’. It shall be highlighted here that statistically sophisticated or complex methods do not necessarily provide more accurate forecasts compared to simple methods (Makridakis & Hibon, 2000). Furthermore, this study is limited by the availability of data. However, the forecasting methods and explanatory variables have been selected systematically.

3.5 Steps involved in forecasting

Most of the technical details of the time series forecasting methods presented here can be referred to the book written by Hanke and Wichern (2005). This book provides a concise step-by-step approach of forecasting methodology. In general, forecasting procedures involve extending the trend of the past into the future. However, the conditions generated from past data may not necessarily occur in the future. Hence, in order to prevent inaccurate forecasts, the assumptions of the models are modified based on the judgement of the researcher. The appropriate sequence of steps in forecasting will help identify indistinguishable data and this helps to minimize errors. According to Hanke and Wichern (2005), forecasting methods operate on data generated by historical events and consist of the following five steps:

- (1) Data collection.
- (2) Data reduction or condensation.
- (3) Model construction and evaluation.

- (4) Model extrapolation (i.e. the actual forecast).
- (5) Forecast evaluation.

It shall be noted that a number of conditions need to be considered during data collection. It is essential that reliable data are used in road safety research and for this reason, the data are retrieved only from authorized publications. However, these data should be interpreted with caution because the number of casualties presented in the reports may be underreported. Previous studies have listed pertinent data on road traffic casualties, which reduces the time taken for data collection in this study. Data reduction (otherwise known as data condensation) is necessary since some of the data may be irrelevant to the problem and may reduce the accuracy of the forecasting models. Even though some of the data may have shown significant correlation with the number and rate of casualties in previous studies, it is possible that this may not be the case in the future.

In this study, it is deemed necessary to develop and evaluate several different models in order to determine the best forecasting model. The forecasting error serves as the basis for model construction and evaluation. In general, simple models are preferable – however, complex models may provide more accurate forecasts since various factors are taken into consideration. The ability to make a wise decision regarding the best forecasting model is indeed an advantage, but it is to some extent based on one's own experience. Extrapolation is conducted to generate data from the observed historical values in order to check the accuracy of models and determine forecasting errors.

Forecast evaluation involves comparing the forecasted values with actual historical values. A variety of forecasting methods have been proposed in previous studies with varying degrees of success. Forecasting errors consist of the sum of the absolute values of errors, average forecast error and the sum of squared errors, which are used to compare the accuracy of a particular method from the alternative forecasting methods considered.

It shall be noted that some forecasting methods also track the magnitude of the error terms over the forecasting period. These error patterns will reveal whether the method is suitable for forecasting, which will help an analyst to select the appropriate forecasting method.

3.6 Descriptive statistical procedures

3.6.1 Numerical summary

The most common descriptive statistical procedure is averaging the data values. The averaging process can be done by adding all of the data values and dividing the sum by the number of values. The result is known as the *mean* (\bar{X}). The mean shows the central tendency of a group of values. Analysts are usually interested in the extent to which the values are dispersed around the mean. The standard deviation is a unit of measurement used to measure the deviation of the values from the mean. The term $n - 1$ is called the degrees of freedom which indicates the number of data items that are independent of one another. The standard deviation represents the sum of squared differences between the measured values and their mean, and is typically extended to the sample *variance*. The variance (S^2) of the sample data is the *standard deviation squared*.

In population statistics, the symbol used for mean, variance and standard deviation differs from those used in sample statistics. The symbol for population mean is μ , whereas the symbol used for population variance and population standard deviation is σ^2 and σ , respectively.

Visualizing data through charts and graphs is usually the first task in data analysis since this task enables analysts to observe the basic features of the data, including unusual observations and unique patterns. Indeed, some variations observed in the data can be inferred from time graphs. Another useful graph in statistical analysis is scatter diagram,

which are used to visualize the relationship between two variables. Time series plots are often used to present chronological data such as the number of road traffic fatalities, in which the number of fatalities is plotted as a function of time.

3.6.2 Probability distributions

A *random variable* is a variable that can contain any value during different trials of an experiment – the exact outcome being a chance or a random event. The random variable is called a *discrete variable* if only certain values are possible. The number of rooms in a hotel is an example of a discrete variable. In contrast, a *continuous variable* is a variable that can contain any value of a random variable within a certain range. The weight of students in a class is an example of a continuous variable.

The *probability distribution* of a discrete random variable is a list of all possible values that a variable can take, along with the probability of each. The expected value of a random variable is the mean value that the variable assumes over a large number of trials. The expected value of a discrete probability distribution can be determined using the following equation:

$$E(X) = \sum [X \cdot P(X)] \quad (3.1)$$

where:

$E(X)$ = Expected value of random variable X

X = Random variable

$P(X)$ = Probability of the random variable X .

A distribution that represents many real-life variables measured on a continuous scale is known as a *normal distribution*. Many useful populations of numbers can be approximated by this distribution. Knowledge on the mean and standard deviation is

required to identify a specific normal distribution. The distribution follows a normal, bell-shaped curve which is symmetrical, as shown in Figure 3.2. The probabilities of the values drawn from a normal distribution that fall within specific limits are first determined by converting these limits into standard deviation units called *Z-scores*. The *Z-score* of any *X* value is the number of standard deviations from the central value of the curve (μ) to that value. The formula used to calculate the *Z-scores* is given by:

$$Z = \frac{X - \mu}{\sigma} \quad (3.2)$$

where:

X = Value of interest

μ = Mean

σ = Standard deviation.

The normal distribution table is constructed using the *Z-score* values, and this table can be used to determine the area under the normal distribution curve between the centre of the curve (μ) and the value of interest *X*. If random variable *X* has a normal distribution, then the *Z-scores* of the random variable have a normal distribution with mean (μ) = 0 and standard deviation (σ) = 1.

3.6.3 Sampling distribution

A sampling distribution is defined as the distribution of all possible values of the sample statistic that can be obtained from the population for a given sample size. Let us consider the following example, whereby a random sample of 100 people is taken from a population. The height of each person in the sample is measured and the mean height is computed. The sample mean (\bar{X}) is assumed as having been drawn from the distribution

of all possible sample means of samples each having a sample size of 100 that can be taken from the population. Similarly, it is assumed that each sample statistic that is computed from the sample data can be considered as having been drawn from a sampling distribution.

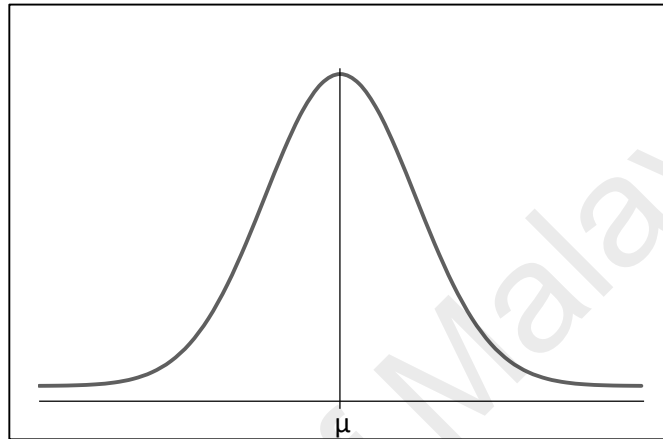


Figure 3.2: Normal distribution curve with population mean μ

According to the central limit theorem, the sampling distribution of the sample means will approximate the normal distribution if the sampling size is sufficiently large. The standard error of the sample mean is given by σ/\sqrt{n} . It shall be noted that the sampling distribution will approximate a normal distribution regardless of the shape of the population distribution from which the samples are drawn. The central limit theorem greatly facilitates analysts in data analysis since it enables one to compute the probability of various sample results based on the probabilities of the normal distribution curve.

However, further analysis is required if the sample size is small. In this case, it is assumed that the population under investigation is normally distributed, and the population standard deviation is unknown and therefore, it must be estimated using the sample standard deviation. In this case, values from the t -distribution are used to

demarcate the sampling distribution area and only the *degrees of freedom* (df) is required. Once the degrees of freedom is known, the t values that exclude the desired percentages of the curve can be determined.

3.6.4 Estimation

There are two types of estimation values which can be determined from a forecasting process: *point estimate* and *interval estimate*. The point estimate of a population parameter is a value calculated from the sample data that estimates the unknown population value. This involves the mean, variance and standard deviation. In contrast, an interval estimate or *confidence interval* is the interval where it is highly likely that the population parameter of interest lies within it. The confidence interval is determined by demarcating an interval around the point estimate and it is generally computed using either the normal distribution or t -distribution.

3.6.5 Hypothesis testing

Hypothesis is an educated guess regarding the population and it is usually a part of statistical analysis. Hypothesis testing is required to determine whether the findings of a study support or reject the hypothesis made at the beginning of the study. The steps involved in hypothesis testing are described briefly as follows:

- (1) Formulate the null hypothesis (H_0) and alternative hypothesis (H_1). The null hypothesis is the outcome of the study that an analyst wants to nullify or refute. In contrast, the alternative hypothesis is one of the possible outcomes expected in the study, and it is also known as the working hypothesis. If H_0 is rejected, then H_1 is accepted.
- (2) Select the significance level (α), which is usually 0.05 or 0.1.

- (3) Calculate the standard error of the sample mean σ/\sqrt{n} . Following this, define the lower and upper limits for rejection of H_0 . This can be done by adding and subtracting the population mean (μ) with/from the product of the t value (at $\alpha = 0.05$ or 0.1) and standard error.
- (4) If the sample mean (\bar{X}) is less or higher than the lower and upper limits for H_0 rejection, reject the null hypothesis (H_0) in favour of the alternative hypothesis (H_t).

3.6.6 Correlation analysis

One of the essential tasks in the development of forecasting models is to examine the relationship between two variables. There are two techniques used for this purpose, namely correlation analysis and regression analysis.

3.6.7 Scatter diagrams

One of the simplest methods used to examine the relationship between two variables is to plot a scatter diagram. Scatter diagrams show whether the dependent variable (plotted on the ordinate or Y-axis of the diagram) tends to increase or decrease with changes in the independent variable (plotted on the abscissa or X-axis of the diagram). Hence, a preliminary conclusion can be drawn regarding the relationship between both of the variables (x and y). There are various types of scatter diagrams – some show whether the relationship between the two variables is strong or weak while others show whether the relationship is positive or negative. Some scatter diagrams will reveal whether the relationship is linear or non-linear while others may indicate that there is no relationship at all between the two variables. One can draw a straight line that passes through most of the points in the scatter diagram in order to determine whether there is a strong linear

relationship between two variables. If most of the data points lie in close proximity to the straight line, then there is a strong linear relationship between the variables. Likewise, if most of the data points lie farther away from the straight line, the linear relationship between the variables is rather weak. The scatter diagram is a good visual representation of the relationship between two variables. However, there are cases whereby more details are required and this can be achieved by measuring the correlation coefficient of the two variables.

3.6.8 Correlation coefficient

The correlation coefficient is a measure of the strength of the linear relationship that exists between the two variables of interest. The correlation coefficient value varies between -1 and +1. If there is a positive relationship between the two variables (such that an increase in x will result in an increase in y) the correlation coefficient will have a positive value. Similarly, if there is a negative relationship between the two variables (such that an increase in x will result in a decrease in y), the correlation coefficient will have a negative value. A correlation coefficient value of +1 and -1 indicates a perfectly linear relationship in the positive and negative direction, respectively. In contrast, the correlation coefficient value is 0 if there is no linear relationship between the two variables. The correlation coefficient is represented by r and is determined using the following formula:

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}} \quad (3.3)$$

3.7 Examination of data patterns and selection of forecasting methods

In general, there are two types of data that are of interest: *cross-sectional* data and *time series* data. Cross-sectional data are obtained from observations of various variables at the same period. This is often useful when carrying out data stratification. However, *time series* data are observations of a variable made over successive increments of time.

3.7.1 Examination of time series data patterns

The time series data patterns are examined in order to select the appropriate forecasting methods. There are various types of time series data patterns: (1) horizontal, (2) trend, (3) seasonal, and (4) cyclical. A horizontal pattern exists if the time series data fluctuate around a constant level or mean. This type of data series is said to be stationary in its mean. A trend pattern exists if the data increase or decrease over an extended period of time. A seasonal pattern exists if the data are influenced by seasonal factors such as weather conditions, holidays, or the length of calendar months. Thus, time series data may have a similar pattern within the same month or within the same quarter of several years. A cyclical pattern exists if the data show rises and falls within different periods. A cyclical pattern is evident if there are wave-like fluctuations in the data series.

3.7.2 Examination of data patterns with autocorrelation analysis

Time series data are often related or correlated. The *autocorrelation coefficient* is measured to determine the correlation between successive time series data. Autocorrelation is the correlation between a variable that is lagged one or more periods and itself. The time series data patterns can be defined based on the autocorrelation coefficient of a variable at different time lags. The autocorrelation coefficient (r_k) between observations Y_t and Y_{t-k} at lag k is computed using the following equation:

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad k = 1, 2, 3, \dots \quad (3.4)$$

where:

r_k = Autocorrelation coefficient for a lag of k periods

\bar{Y} = Mean of the values of the data series

Y_t = Observation in time period t

Y_{t-k} = Observation at k time periods earlier or at time period $t - k$.

The various values of the autocorrelation coefficient at lag k (r_k) are typically presented in the form of a graph, which assists analysts in analysing the data patterns. This graph is known as *autocorrelation function* or *correlogram*. The values of subsequent autocorrelation coefficients will show whether the time series data are *random*, *non-stationary*, *stationary* or *seasonal*. If a series is random, the autocorrelation coefficients between Y_t and Y_{t-k} for any lag k are close to zero. If a series has a trend, the successive observations will be highly correlated, and the autocorrelation coefficients typically differ significantly from zero for the first several time lags and then decrease to zero as the number of lags increases. The autocorrelation coefficient for time lag 1 is often very large (i.e. close to 1). The autocorrelation coefficient for time lag 2 is also large though it is not as large as that for time lag 1. If a series has a seasonal pattern, a significant autocorrelation coefficient will occur at the seasonal time lag or at multiples of the seasonal lag. The seasonal lag is 4 and 12 for quarterly data and monthly data, respectively.

The *portmanteau* test is used to determine whether the first several autocorrelation coefficients of a time series data are significantly different from zero. The Ljung-Box test is one of the common portmanteau tests, which involves evaluating the Q test statistic.

This test is usually applied to the residuals of a forecast model. If the autocorrelations are computed from a *random (white noise)* process, the statistic Q has a *chi-square* distribution with m degrees of freedom (the number of time lags to be tested). For the residuals of a forecast model, however, the Q test statistic has a chi-square distribution in which the degrees of freedom equal to m minus the number of parameters estimated in the model. The value of Q can be compared with the value in the chi-square distribution table to determine if it is larger than its expected value under the null hypothesis (all autocorrelation coefficients in the set are zero). Alternatively, the p value generated by Q can be computed and interpreted. The Q test statistic is given by:

$$Q = n(n + 2) \sum_{k=1}^m \frac{r_k^2}{n - k} \quad (3.5)$$

where:

n = Number of observations in the time series

k = Time lag

m = Number of time lags to be tested

r_k = Sample autocorrelation function of the residuals lagged at k time periods.

A stationary (white noise) time series is one whose basic statistical properties, such as the mean and variance, remain constant over time. Consequently, a series that varies about a fixed level over time (i.e. no growth or decline) is said to be stationary. In contrast, a series that contains a trend is said to be non-stationary. The autocorrelation coefficients for a stationary series will decrease to zero fairly rapidly – generally after the second or third time lag. The sample autocorrelations remain fairly large for non-stationary series over several time periods. A common practice in analysing a non-stationary series is to remove the trend before further modelling. This involves transforming the series by using

logarithms, square roots or differences. The differencing method is often used to remove the trend from a non-stationary series.

According to Hanke and Wichern (2005), the methods which need to be considered when forecasting time series data are listed as follows:

- (1) Methods for stationary series: naïve methods, simple averaging methods, moving averages, and autoregressive moving average (ARMA) models (Box-Jenkins methods).
- (2) Methods for trending series: moving averages, Holt's linear exponential smoothing, simple regression, growth curves, exponential models, and autoregressive moving average (ARMA) models (Box-Jenkins methods).
- (3) Methods for cyclical series: classical decomposition, econometric models, multiple regression, and ARIMA models (Box-Jenkins methods).

3.7.3 Forecasting errors

It is essential that the forecasting error is measured in order to determine the accuracy of a forecasting method. The error of a forecasting method can be determined by measuring the differences between the observed values and forecast values. These differences are termed as *residuals*. Makridakis and Hibon (2000) measured the accuracy of 24 methods used to forecast 3003 time series data. They implemented five accuracy measures (symmetric MAPE, average ranking, median symmetric APE, percentage better and median RAE) to analyse the performance of the various forecasting methods. The mean average percentage error (MAPE) is chosen in this study, following the recommendations of Hyndman and Koehler (2006). If all of the series are on the same scale, then the mean absolute error (MAE) is preferable because it is simpler to explain. However, if all of the data are positive and much greater than zero, the MAPE is

preferable for simplicity. However, in circumstances where there are differing scales, including data that are close to zero or negative, Hyndman and Koehler (2006) suggested that the mean absolute scaled error (MASE) is the best measure of forecast accuracy.

The MAPE represents the forecasting error in terms of percentage rather than the absolute value. It is computed by determining the absolute error for each period. The absolute error is then divided by the actual observed value for each period. Finally, the absolute percentage errors are averaged. This approach is useful when the size or magnitude of the forecast variable is important in evaluating the accuracy of the forecast. The MAPE provides an indication of how large the forecast errors are in comparison to the actual values of the data series, and is especially useful when the Y_t values are large. The MAPE can also be used to compare the accuracy of similar or different forecasting methods on two entirely different series. The MAPE is computed using the following equation:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - \hat{Y}_t|}{Y_t} \quad (3.6)$$

3.8 Components of a time series

3.8.1 Decomposition

Various variables often contribute to time series data. Hence, it is imperative to identify the component factors that influence each of the values in a data series. This identification procedure is known as decomposition. The component factors are: (1) trend, (2) cyclical, (3) seasonal, and (4) irregular or random. Identification of each factor is carried out separately. Once the component factors of a time series are identified, the next step is to forecast the values of the time series according to each component factor.

The combination of values attained from each forecast yields the final forecasting values of the time series. A time series data is typically fit for one of the component factors. This speeds up the process since some of the component factors can be eliminated. The decomposition is used to gain understanding on the time series patterns as more flexible forecast modelling methods are made available in recent years.

The trend component will cause the time series to either increase or decrease. In contrast, the cyclical component will cause the series to show wavelike fluctuations or cycles of more than a duration of one year. The cycles are usually difficult to identify and they cannot be separated from the trend. Seasonal fluctuations are typically found in quarterly, monthly or weekly data. Seasonal variation repeats itself year after year. Seasonal patterns occur due to the influence of the weather, as well as national and school holidays. The *irregular component* consists of unpredictable or random fluctuations resulting from a variety of events. Even though an individual event may not produce a marked effect, the combination of these events may result in a significant effect.

The relationship between these components and the time series can be described by two types of mathematical models. In the first model, the time series values are treated as the sum of the components, and hence it is known as an *additive* model. In the second model, the time series values are treated as the product of the components, and therefore it is known as a *multiplicative* model. According to Yannis et al. (2011a), a number of researchers presumed that the explanatory variables have a multiplicative effect on road traffic accidents rather than an additive effect. The mathematical representation for the additive and multiplicative models is given by Equation (3.7) and Equation (3.8), respectively.

$$Y_t = T_t + S_t + I_t \quad (3.7)$$

$$Y_t = T_t \times S_t \times I_t \quad (3.8)$$

where:

Y_t = Observed value

T_t = Trend component

S_t = Seasonal component

I_t = Irregular component.

The additive model is suitable when the time series being analysed has roughly the same variability throughout the period of the series. In this model, all of the values of the series fall within a band of constant width centred on the trend. However, for the multiplicative model, the variability of the time series increases with the level. The values of the series disperse as the trend increases.

The trend of a time series may be described by a *straight line* or a *smooth curve*. However, it is rather rare for a time series to be represented by a straight line, as given in Equation 3.9. A straight line is usually used to define the general direction of the observed time series. On the other hand, a large number of time series can be represented by a smooth curve. The life cycle of most time series consists of three stages: introduction, growth and saturation. The *quadratic* curve is the simplest smooth curve that provides good fit for a time series. Other forms of smooth curves include *exponential*, *Gompertz* and *logistic* curves. The mathematical representation of a smooth curve that follows a quadratic and exponential trend is given by Equation (3.10) and Equation (3.11), respectively.

$$\hat{T}_t = a_0 + a_1 t \quad (3.9)$$

$$\hat{T}_t = a_0 + a_1 t + a_2 t^2 \quad (3.10)$$

$$\hat{T}_t = a_0 a_1^t \quad (3.11)$$

The use of an exponential trend may sometimes overestimate the results. This is one of the problems with the exponential trend model. This problem may be eliminated by applying the Gompertz and logistic trend. Even though the features of logistic and Gompertz trends are very similar, the logistic curve trend has a slightly gentler slope. Both models are appropriate to forecast series that begins with a low value, followed by an increase and finally reaches a saturation point.

3.8.2 Forecasting trend

The trend of a time series can be observed from a time series plot which shows the changes in the value of a variable as a function of time. The trend may either be linear, quadratic, exponential, Gompertz or logistic. Even though the type of trend can be identified based on the time series plot, this approach may be less accurate, particularly if the trend approximates the pattern of a smooth curve. Hence, two objective criteria are used to provide a more accurate means to select and use a trend curve, namely Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). BIC is also known as Schwarz's Bayesian Criterion (SBC). Both criteria have been proven to be very useful in determining the most appropriate time series model. Both AIC and SBC will be described in detail in this chapter.

3.8.3 Seasonality

A time series that repeats itself year after year shows a *seasonal pattern*. Only data observations that are conducted weekly, monthly, or quarterly will show seasonality. Seasonal patterns cannot be used for annual data which are recorded only once a year.

This pattern is useful only for short-term forecasting. It has been mentioned previously that the data pertinent to road traffic safety are published annually and therefore, it can be expected that the data series will not reflect a seasonal pattern. The methods used to measure seasonal variations such as cyclical and irregular variations are not elaborated here.

3.9 Regression analysis

There are two types of regression analysis: simple linear and multiple linear regression analysis. Multiple linear regression is a direct generalization of simple linear regression. The primary difference between these statistical methods is that the value of the dependent variable is dependent on more than one independent variable in multiple regression analysis. In road traffic safety research, the relationship between the number and/or rate of casualties with the explanatory variables is not that straightforward. In most cases, it is necessary to use more than one independent variable to predict a dependent variable accurately. For this reason, the multiple regression model is suitable to determine the relationship between the dependent and independent variables in this study. The relationship between the dependent variable (Y) and independent variable (X) of a multiple regression model is given by:

$$Y = b_0 + b_1X_1 + b_2X_2 + \cdots + b_kX_k + \varepsilon \quad (3.12)$$

The multiple regression model includes the *error components* (ε) where $\varepsilon \sim N(0, \sigma^2)$, which are also known as residuals (e). The residuals represent the deviations of the response from the true relationship. The errors are basically unobservable random variables which account for the effects of other factors on the response. In linear and multiple regression analyses, the errors are assumed to be independent and each error is normally distributed with a mean of 0 and an unknown standard deviation (σ). The

regression coefficients (b_0, b_1, \dots, b_k) can be estimated using the least squares method. The first coefficient, b_0 , represents the Y-intercept. In the least squares method, the values of the regression coefficients are chosen such that the sum of squared errors (distances between the actual Y and estimated \hat{Y}) is minimized. The sum of squared errors (SSE) is given by:

$$SSE = \sum_i (Y_i - b_0 - b_1X_{i1} - b_2X_{i2} - \dots - b_kX_{ik})^2 = \sum_i (Y - \hat{Y})^2 \quad (3.13)$$

where:

SSE = Sum of squared errors

Y = Actual values

\hat{Y} = Estimated values

i = 1, 2, 3, ..., n

k = Number of independent variables in the regression analysis.

3.9.1 Inference for the regression models

In regression analysis, there is a need to measure the extent to which the sample data points are spread around the fitted regression function. This is called the *standard error of the estimate* and its objective is to measure the amount by which the actual values differ from the estimated values. The standard error of the estimate is similar to the standard deviation of the residuals and it is determined using the following equation:

$$S_{y,x} = \frac{\sqrt{\sum (Y - \hat{Y})^2}}{n - k - 1} \quad (3.14)$$

where:

$S_{y,x}$ = Standard error of the estimate

Y = Actual values

\hat{Y} = Estimated values

n = Number of data

k = Number of independent variables in the regression.

Another inference of the regression model is to define the significance of the regression. This can be obtained by measuring the *multiple correlation coefficient* (R) of the model, which shows the correlation between the responses (Y) and fitted values (\hat{Y}). R is the square root of the coefficient of determination (R^2), which is given by:

$$R^2 = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \bar{Y})^2} \quad (3.15)$$

Thus, R falls within a range of $0 \leq R \leq 1$.

3.9.2 Multi-collinearity

In most multiple regression analyses, the data are routinely recorded rather than generated from pre-selected settings of the independent variables. The independent variables are frequently linearly dependent. In other words, some of the independent variables always move together, whether they are increasing or decreasing. In this case, even though the regression model coefficients can be obtained, these estimates tend to be unstable such that the values can change dramatically with slight changes in the data and the values are larger than expected. The t statistic used to judge the significance of individual terms may be insignificant. However, the F test will indicate that the regression is significant. Hence, the calculation of the least squares estimates is sensitive to rounding errors.

The linear relationship between two or more independent variables is termed as *multi-collinearity*. The multi-collinearity between variables is measured by the *variance inflation factor* (VIF) which is given by:

$$VIF_j = \frac{1}{1 - R_j^2} \quad j = 1, 2, 3, \dots k \quad (3.16)$$

where:

R_j^2 = Coefficient of determination from the regression of the j^{th} independent variable on the remaining $k - 1$ independent variables

k = Number of independent variables.

A *VIF* value close to 1 indicates that multi-collinearity is not a problem for that independent variable. Its estimated coefficient and associated t value will not change significantly as the other independent variables are added or deleted from the regression equation. A *VIF* much greater than 1 indicates that the estimated coefficient associated with that independent variable is unstable. If the *VIF* is relatively high, the t statistic may change considerably when the other independent variables are added or deleted from the regression equation. Statisticians recommend that the value of *VIF* should be less than 10 in order to prevent collinearity problems.

3.9.3 Stepwise regression

Various explanatory variables can be included in a multiple regression model. Nevertheless, a model with a higher number of variables is not necessarily the best one. In fact, multi-collinearity issues are often present in a complex model. Consequently, it is crucial for one to select the best regression model. The following steps are used to select the best regression model.

- Step 1. Include all of the potential predictor variables.
- Step 2. Remove the independent variables that seem inappropriate for the regression model. Several criteria should be taken into account, but priority should be given to remove independent variables that duplicate other independent variables (multi-collinearity).

There are two approaches used to select the best regression model: (1) all possible regressions and (2) stepwise regression. The latter approach is adopted in this study. Stepwise regression is useful to simplify the regression model. In this approach, the variables are added or removed from the regression model one at a time in order to attain a model that contains only significant predictors but at the same time, it does not exclude any useful variables. The basic steps involved in stepwise regression are listed below:

- Step 1. Obtain all of the possible simple regressions. The predictor variable that has the largest correlation with the independent variable is the first variable which will be included in the regression equation.
- Step 2. Include the variable that gives the largest significant contribution to the regression sum of squares as the next variable in the regression equation. Determine the significance of the contribution using F test. The value of the F statistic that needs to be exceeded before the contribution of a variable is deemed significant is known as F to enter.
- Step 3. Check the significance of the new equation using F test. If the F statistic is less than the F to remove, delete the variable from the regression equation.
- Step 4. Repeat Steps 2 and 3 until all possible additions are non-significant and all possible deletions are significant. The selection process ends at this point.

Stepwise regression is carried out using Statgraphics. The software provides two stepwise options: *forward* and *backward selection*. Forward selection begins with the model that contains only one constant and brings the variables in one at a time if they improve the fit significantly. In contrast, backward selection begins with the model that contains all of the variables and removes them one at a time until all of the remaining variables are statistically significant. Backward selection is adopted in this study.

3.10 Regression of time series data

Many applications of forecasting, including forecasting the number and rate of road traffic fatalities, are associated with time series data. Since the data collected over time tend to exhibit either one of the following (trend, seasonal, cyclical or irregular patterns), the observations at different time periods will be related to one another or in other words, autocorrelated. Hence, the sample of observations on time series data cannot be regarded as a random sample, stationary (white noise) or non-stationary data. Hence, insight analysis is required when fitting regression models to time series data.

3.10.1 The autocorrelation problem

The relationship between a dependent variable (Y) and an independent variable (X) of a linear regression model is given by:

$$Y = b_0 + b_1X + \varepsilon \quad (3.17)$$

The statistical model includes the *error components* (ε) that represent the deviations of the response from the true relationship. The errors are unobservable random variables accounting for the effects of other factors on the response. In linear and multiple

regression, the errors are assumed to be independent and each error is normally distributed with a mean of 0 and an unknown standard deviation (σ). Conversely, in most time series data analyses, it is assumed that the errors (ε) are not independent of one another. This is due to the fact that the value of a variable in the current year may be related or correlated with the value in the previous year or the value of a variable several years ago. Since the values of the variables in different years are not independent, they are autocorrelated. Autocorrelation occurs when the effect of an independent variable on the response variable is distributed over time. The first-order *serial correlation* of a simple linear regression model is given by Equation (3.17). In this correlation, the error term in the current time period is directly related to the error term in the previous time period which is given by the following equation:

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t \quad (3.18)$$

where:

ε_t = Error at time t

ρ = Parameter (lag 1 autocorrelation coefficient) that measures the correlation between adjacent error terms

v_t = Normally distributed independent error term with zero mean and variance ($v_t \sim N(0, \sigma^2)$).

It is evident that the level of one error term (ε_{t-1}) directly affects the level of the next error term (ε_t). The magnitude of the autocorrelation coefficient ρ , where $-1 < \rho < 1$, indicates the strength of the serial correlation. If ρ is zero, then there is no serial correlation, and the error terms are independent ($\varepsilon_t = v_t$).

The problem arises when there is a strong autocorrelation between the variables since this will make two unrelated variables appear to be related. Hence, the application of standard regression procedures to the observed variables may result in a significant regression and consequently, the estimated relationship will be spurious. Examination of the residuals will provide insight on the problem. It shall be noted that if standard regression procedures are applied without caution, any spurious regression may go undetected, which leads to misinterpretation of the results. When the time series data are autocorrelated, the residuals need to be examined in the regression model. Otherwise, the conclusions drawn are not justified. Several assumptions need to be made when solving autocorrelated time series data. This is due to the fact that the standard error of the estimate can seriously underestimate the variability in the error terms and the usual inferences based on the t and F statistics will no longer be applicable. This in turn, results in spurious regressions.

3.10.2 Test for serial correlation

The two types of analysis used for serial correlation are highlighted in this section. The first method involves evaluating the Q test statistic developed by Ljung and Box, as described in Section 3.6.2. The second method involves carrying out the *Augmented Dickey-Fuller* (ADF) test. In this test, it is assumed that the non-stationary condition is attributed to unit root problems. Based on the literature related to time series, random walk models are also known as unit root processes. In the simplest case of an uncorrelated error term, the test is initiated by estimating the following equation:

$$Y_t = \rho Y_{t-1} + e_t \quad (3.19)$$

By taking the first difference, Equation (3.19) can be written as:

$$\Delta Y_t = Y_t - Y_{t-1} = (\rho - 1)Y_{t-1} + e_t = \delta Y_{t-1} + e_t \quad (3.20)$$

The null hypothesis is $\delta = 0$, which means that $\rho = 1$. If the null hypothesis cannot be rejected, the data generation process is inferred as having a unit root and is therefore, non-stationary. This implies that the time series under consideration is integrated. The creators of this test, Dickey and Fuller, have shown that under the null hypothesis, the estimated t value of δ follows the τ statistic, and the critical τ values are computed using Monte Carlo simulations. They also introduced a competing F test with the usual F computations – however, with special critical values. The series may be non-stationary because of random walk, drift or trend.

- (1) Y_t is a random walk whereby $\Delta Y_t = \delta Y_{t-1} + e_t$.
- (2) Y_t is a random walk with drift $\Delta Y_t = \alpha + \delta Y_{t-1} + e_t$.
- (3) Y_t is a random walk with drift and trend $\Delta Y_t = \alpha + \beta t + \delta Y_{t-1} + e_t$.

In each case, the null hypothesis is $\delta = 0$. The alternative hypothesis is $\delta < 0$ (one-sided test) or in other words, the time series is stationary. Either one of the conclusions can be drawn if the null hypothesis is rejected in favour of the alternative hypothesis: (1) Y_t is a stationary series with a zero mean, (2) Y_t is stationary with a non-zero mean or (3) Y_t is stationary around a deterministic trend. The SAS software is used to perform the ADF test. In this study, the null hypothesis is given by the following statement: the response series is non-stationary, whereas the alternative hypothesis is given by the following statement: the response series is stationary. If the probability is found to be greater or equal to 0.05, then the null hypothesis cannot be rejected and thus, it is concluded that the series is non-stationary. Hence, transformation needs to be carried out to make the data series stationary.

3.10.3 Solutions for autocorrelation problems

In general, a time series must be stationary in variance and mean before an analysis is carried out. However, if the variables are autocorrelated in any way, the series will be non-stationary. In order to make a series stationary, either one of the following techniques can be used such as log transformation, power transformation, or Box–Cox transformation techniques. However, differencing techniques are used to ensure that the series is stationary in mean.

3.10.3.1 Box-Cox transformation

Box-Cox transformation is the most commonly used variance-stabilizing transformation technique. This transformation changes the variance of the residuals into a constant. The Box-Cox transformation is given by:

$$y = \begin{cases} \frac{(Y^\lambda - 1)}{\lambda}, & \lambda \neq 0 \\ \text{Log } Y, & \lambda = 0 \end{cases} \quad (3.21)$$

where λ is the shape parameter and a real number. The various types of Box-Cox transformation based on the value of λ are presented in Table 3.2. In this study, Minitab software is used to determine the suitable type of Box-Cox transformation.

Table 3.2: Box-Cox transformation based on the value of λ

| λ | Transformation |
|-----------|-------------------------------|
| 2.0 | $y = Y^2$ |
| 0.5 | $y = \sqrt{Y}$ |
| 0.0 | $y = \log(Y)$ or $y = \ln(Y)$ |
| -0.5 | $y = \frac{1}{\sqrt{Y}}$ |
| -1.0 | $y = \frac{1}{Y}$ |
| -2.0 | $y = \frac{1}{Y^2}$ |

3.10.3.2 Regression with differencing

When there is high autocorrelation in the data, the model itself changes rather than the levels, which often eliminates serial correlation. In this study, differencing is considered when the probability obtained from the ADF test is greater or equal to 0.1.

A simple regression model for time period t and $t-1$ is given by:

$$Y_t = b_0 + b_1 X_t + \varepsilon_t \quad (3.22)$$

$$Y_{t-1} = b_0 + b_1 X_{t-1} + \varepsilon_{t-1} \quad (3.23)$$

with

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t$$

where:

ρ = Correlation between consecutive errors

v_t = Random error

v_t = ε_t when $\rho = 0$.

In addition, the general equation used to compute the residual autocorrelation coefficient for a lag of k periods is given by:

$$r_{k(\varepsilon)} = \frac{\sum_{t=k+1}^n \varepsilon_t \varepsilon_{t-1}}{\sum_{t=1}^n \varepsilon_t^2} \quad k = 1, 2, 3, \dots, K \quad (3.24)$$

where:

$r_{k(\varepsilon)}$ = Residual autocorrelation coefficient for a lag of k periods

n = Number of residuals

K = The maximum of k periods, typically $n/4$.

Multiplying both sides of Equation (3.23) by ρ gives:

$$\rho Y_{t-1} = \rho b_0 + \rho b_1 X_{t-1} + \rho \varepsilon_{t-1} \quad (3.25)$$

Following this, subtracting Equation 3.22 from Equation 3.25

$$(Y_t = b_0 + b_1 X_t + \varepsilon_t) - (\rho Y_{t-1} = \rho b_0 + \rho b_1 X_{t-1} + \rho \varepsilon_{t-1})$$

gives:

$$Y_t - \rho Y_{t-1} = b_0 - \rho b_0 + (b_1 X_t - \rho b_1 X_{t-1}) + (\varepsilon_t - \rho \varepsilon_{t-1}) \quad (3.26)$$

which can be simplified as:

$$Y'_t = b_0(1 - \rho) + b_1 X'_t + v_t \quad (3.27)$$

where the generalized differences are:

$$Y'_t = Y_t - \rho Y_{t-1}$$

$$X'_t = X_t - \rho X_{t-1}.$$

The errors v_t are independently distributed with a mean equal to zero and a constant variance. Thus, the usual regression methods can be applied to the regression model. However, if the correlation between consecutive errors is strong (ρ close to 1), then the

generalized differences are essentially simple or the first differences, and the intercept term in the model which is expressed by $(b_0(1 - \rho))$ is close to zero (disappears). The regression models constructed using generalized differences can be used to eliminate the serial correlation of the time series data.

$$Y'_t = Y_t - Y_{t-1}$$

$$X'_t = X_t - X_{t-1}.$$

3.10.3.3 Autoregressive models

Autocorrelation occurs when the values of the dependent variable in one time period are linearly related to their values in another time period. Autocorrelation can be eliminated by modelling directly the association between different time periods. This is done by using the dependent variables lagged one or more time periods as the predictor or independent variables, which results in an autoregressive model. The first-order autoregressive model is written as follows:

$$Y_t = b_0 + b_1 Y_{t-1} + \varepsilon_t \quad (3.28)$$

where the errors (ε_t) are assumed to have the usual regression model properties. Once this model has been fitted to the data by least squares, the forecasting equation becomes:

$$\hat{Y}_t = b_0 + b_1 Y_{t-1} \quad (3.29)$$

When regression analysis is applied to the time series data, the residuals are frequently autocorrelated (serial correlation). Problems may arise since errors are assumed to be

independent in a regression analysis. The R^2 can be artificially high for a regression with data containing serial correlation. The standard errors of the regression coefficients can be seriously underestimated and the corresponding t statistics will be inflated.

The autocorrelated residuals may be due to the omission or addition of one or more key predictor variables which in turn, affects the variations of the dependent variable. This problem can be solved by identifying the missing variable(s) to be included in the model or applying regression models with differencing or autoregressive models.

3.10.4 The heteroscedasticity problem

The variability of some time series tends to increase with the level of the series. The variability of a time series can increase if a variable is growing at a constant rate. This non-constant variability is termed as *heteroscedasticity*. In a regression, heteroscedasticity occurs if the variance of the error term (ε) is not constant. If the variability for recent time periods is larger than its value for past time periods, then the standard error of the estimate ($S_{y,x}$) underestimates the current standard deviation of the error term. If the standard deviation of the estimate is then used to set forecast limits for future observations, these limits can be too narrow for the stated confidence level. The problem of heteroscedasticity can be solved by means of simple transformations of the data. For example, in the case of two variables, the log linear model may be used to reduce heteroscedasticity.

3.11 Box-Jenkins (ARIMA) models

Several approaches can be used to forecast a time series such as exponential smoothing, decomposition into trend, seasonal and irregular components, regression models, ARIMA models including ARIMAX, as well as state-space models. These

approaches can be classified into univariate and multivariate analyses. The regression and ARIMAX models are appropriate for time series data. The forecasts of the dependent variable (usually denoted as Y) generated from these models are obtained by forecasting the future values of the independent variables (X_i).

Autoregressive integrated moving average (ARIMA) is a model that can produce accurate forecasts based on the historical patterns of the time series data. ARIMA models are capable of representing stationary time series. The data will be differenced and/or transferred if the time series data contains non-stationary processes (i.e. no constant mean level). In general, stationary processes vary about a fixed level and thus they will have a constant mean. It shall be noted that independent variables are not required in the construction of ARIMA models because these models make use of the information in the series itself to generate forecasts. Conversely, ARIMAX include exogenous (explanatory) variables and thus the forecasts will be influenced by the future values of the variables.

ARIMA models are highly dependent upon the autocorrelation patterns of the data. Box and Jenkins (1976) have advanced the methodology for identifying, fitting and checking the appropriate ARIMA models. For this reason, ARIMA modelling and forecasting is often referred to as the Box-Jenkins methodology.

3.11.1 Box-Jenkins (ARIMA) methodology

The Box-Jenkins methodology of forecasting does not assume a particular pattern in the historical data of the series which will be forecasted. The iterative approach is used to identify a possible model from a general class of models. Following this, the selected model is checked against the historical data to determine whether it accurately describes the series. The model fits well if the residuals are generally small, randomly distributed and provide useful information. If the model generated is unsatisfactory, the process is

repeated using new model parameters in order to improve the previous model. This iterative procedure continues until a satisfactory model is attained. The final model is used for forecasting. Figure 3.3 shows the iterative approach used to construct a Box-Jenkins model for forecasting or control.

The selection process of an ARIMA model is initiated by plotting the time series over the period of interest. The plot is then examined in order to identify its general characteristics. This will enable the analyst to determine whether the series can be classified as having a trend or whether it is seasonal, cyclical or irregular. The next task involves calculating and plotting its autocorrelations for several time lags. At present, there is an abundant of literature which provide various patterns of sample autocorrelations. This greatly facilitates the analyst in matching the calculated autocorrelation of a time series with the known pattern associated with a particular ARIMA model. This matching is done for both autocorrelation function (ACF) and partial autocorrelation function (PACF). The purpose of matching the calculated autocorrelations with the known patterns is to shorten the identification process. The best model is determined by fitting and checking the models based on the AIC and SBC. In summary, the sequence of steps involved in the Box-Jenkins methodology involved identifying, fitting, and checking the time series data.

Advancements in software technology have made it easier to develop time series models based on the Box-Jenkins methodology. Often, the autocorrelations calculated from the data will not exactly match with any set of theoretical autocorrelations linked with a particular ARIMA model. It is known that sampling variations will influence the autocorrelations calculated. For this reason, computer software is extremely useful.

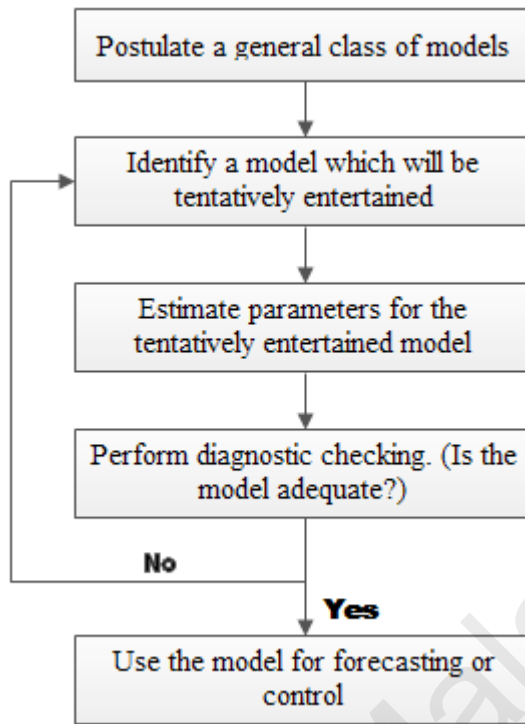


Figure 3.3: Steps involved in the iterative approach to construct a Box-Jenkins model for forecasting or control
Source: Box et al. (1994)

3.11.1.1 Autoregressive models

The mathematical representation of a p^{th} order of autoregressive model is given by:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3.30)$$

where:

Y_t = Response (dependent) variable at time t

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ = Response variable at time lags $t - 1, t - 2, \dots, t - p$, respectively.

The Y 's are all independent variables

$\phi_0, \phi_1, \phi_2, \dots, \phi_p$ = Coefficients to be estimated

ε_t = Error term at time t that represents the effects of variables which are not explained by the model. The assumptions about the error term are the same as those for the standard regression model.

The model given by Equation (3.30) is known as an autoregressive model. This model has the appearance of a regression model with lagged values of the dependent variable in the independent variable positions. Autoregressive models are appropriate for stationary time series and the coefficient ϕ_0 is related to the constant level of theories. The coefficient ϕ_0 is not required if the time series data vary about zero or are expressed as deviations from the mean $Y_t - \hat{Y}$. In a stationary time series, the autocorrelation coefficients will often trail off to zero, whereas the partial autocorrelation coefficients will decrease to zero after the second time lag. Nevertheless, the sample autocorrelation functions will differ from the theoretical functions due to sampling variations.

The forecasts of autoregressive models generally depend on the observed values in previous time periods. For AR(1) models, the forecasts of the next value depend on the observations of a previous time period. For AR(2) models, the forecasts of the next value depend on the observations for two previous time periods, and so forth.

3.11.1.2 Moving average models

The mathematical representation of a q -th order moving average model is given by:

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \cdots - \omega_q \varepsilon_{t-q} \quad (3.31)$$

where:

- Y_t = Response (dependent) variable at time t
- μ = Constant mean of the process
- $\omega_1, \omega_2, \dots, \omega_q$ = Coefficients to be estimated
- ε_t = Error term that represents the effects of variables which are not explained by the model. The assumptions about the error term are the same as those for the standard regression model

$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ = Errors in previous time periods that are incorporated in response Y_t at the time.

It can be seen from both autoregressive and moving average models that the dependent variable Y_t depends on previous values of the errors rather than on the variable itself. The Y_t forecasted by moving average (MA) models is based on a linear combination of a finite number of past errors. In contrast, the Y_t forecasted by autoregressive (AR) is a linear function of a finite number of past values of Y_t . Moving average refers to the deviation of the response from its mean, $Y_t - \mu$, and it is a linear combination of current and past errors. As the time moves forward, the errors involved in the linear combination will move forward.

$$Y_t - \mu = \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q} \quad (3.32)$$

$$Y_{t+1} - \mu = \varepsilon_{t+1} - \omega_1 \varepsilon_t - \omega_2 \varepsilon_{t-1} - \dots - \omega_q \varepsilon_{t-q+1} \quad (3.33)$$

It shall be noted that the weights $\omega_1, \omega_2, \dots, \omega_q$ in Equations (3.32) and (3.33) do not necessarily give a total of to 1 and may be either positive or negative, even though each of them is preceded by a minus sign in the specification of the model.

In moving average models, the coefficients are obtained iteratively using a non-linear least squares algorithm. The aim of the iteration is to achieve the smallest sum of squared errors. A non-linear least squares procedure is often used to estimate the coefficients rather than fitting them using standard regression packages. Following this, the errors corresponding to the time periods that have already occurred are replaced by the residuals for those time periods in order to compute the forecasts. The number of residuals involved in the forecast of the next observation is equal to the order (q) of the moving average model.

3.11.1.3 Autoregressive moving average models

A combination of the autoregressive and moving average models is called the autoregressive-moving average model. The model is denoted by $\text{ARMA}(p,q)$, where p is the order of the autoregressive and q is the order of the moving average. The $\text{ARMA}(p,q)$ model has the following generic mathematic representation:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \cdots - \omega_q \varepsilon_{t-q} \quad (3.34)$$

It shall be highlighted that the time series data must be stationary in order to apply an $\text{ARMA}(p, q)$ model. When $q = 0$, the model changes into a pure autoregressive model of order p (i.e. $\text{ARMA}(p, 0)$). Likewise, when $p = 0$, the model changes into a pure moving average model of order q (i.e. $\text{ARMA}(0, q)$). The forecasts of an $\text{ARMA}(p, q)$ model are dependent upon the current and past values of the response Y as well as the current and past values of the errors (residuals).

The ACF and PACF properties for autoregressive-moving average processes are summarized in Table 3.3. Some examples of ACFs and PACFs of common ARMA models are shown in Figure 3.4.

Table 3.3: Properties of autoregressive (AR), moving average (MA) and mixed autoregressive moving average (ARMA) processes

| Model | Autocorrelations (ACF) | Partial autocorrelations (PACF) |
|--------------|--|--|
| $AR(p)$ | Die out | Cut off after order p of the process |
| $MA(q)$ | Cut off after order q of the process | Die out |
| $ARMA(p, q)$ | Die out | Die out |

The orders p and q in an ARMA model are determined from the patterns of the sample ACF and PACF.

3.11.1.4 Model construction strategy

The steps involved in the Box-Jenkins (ARIMA) approach consist of initial identification of the model, estimation of coefficients and model checking (residual analysis). An iterative process is conducted until the model residuals indicate that no further modification is necessary. The fitted model is then used for forecasting. In this study, SAS software is used to execute the ARIMA models. The steps involved in the Box-Jenkins methodology are detailed as follows.

Step 1: Identification of model

Since the time series data must be stationary (i.e. varies about a fixed level), the first step is to determine whether the series is stationary. To an experienced analyst, the plot of the series along with the sample autocorrelation function may be used to determine whether the time series is stationary. As a rough guideline, a time series data is non-

stationary if the series appears to increase or decrease with respect to time and the sample autocorrelations fail to die out rapidly. Figure 3.5 shows an example of a non-stationary time series, including the patterns of the sample autocorrelations. The data series shown here represents the number of fatalities in Malaysia from 1981 to 2012. It can be observed that the ACF decays slowly even though the PACF dies out. It shall be noted that a large number of time series are non-stationary – however, they can be converted into a stationary time series. The differencing method is usually used for this purpose. Thus, the original series is replaced by a series of differences. An ARMA model is then defined for the differenced series. Consequently, the changes in the model are taken rather than its levels.

The original time series with order $p = 1$ and $q = 1$ has the following mathematical representation. The constant term (ϕ_0) may not be required in an ARMA model with differencing.

$$Y_t - Y_{t-1} = \phi_1(Y_{t-1} - Y_{t-2}) + \varepsilon_t - \omega_1\varepsilon_{t-1} \quad (3.35)$$

If $Y_t - Y_{t-1} = \Delta Y_t$ and $Y_{t-1} - Y_{t-2} = \Delta Y_{t-1}$, the equation above can be written as:

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \varepsilon_t - \omega_1 \varepsilon_{t-1} \quad (3.36)$$

In some cases, first-order differencing is inadequate to yield a stationary data series. Hence, second-order differencing is used for this purpose. This is achieved using the following equation:

$$\Delta^2 Y_t = \Delta(\Delta Y_t) = \Delta(Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2} \quad (3.37)$$

Most of the time, second-order differencing will yield a stationary data series, which is evident when the data varies about a fixed level and the sample autocorrelations die out fairly rapidly. The level of differencing (i.e. 1, 2 ...) to achieve stationarity is denoted by d , and the non-stationary model is denoted by ARIMA(p, d, q). The primary purpose of

the differencing method is to transform a non-stationary series into a stationary series. It shall be noted that the forecasts for the original data series can always be computed directly from the fitted model.

Once a stationary data series is attained, the next step involves computing and plotting the autocorrelations and partial autocorrelations. This step is essential in order to compare the computed autocorrelations and partial autocorrelations with the theoretical ones for various ARIMA models. The theoretical correlations of common ARIMA models can be referred to the literature, and some examples are shown in Figure 3.4. In order to simplify the process, both of the sample autocorrelations and sample partial autocorrelations are compared with $\pm \frac{2}{\sqrt{n}}$, where n is the number of observations in the time series. This process yields the tentative ARIMA models. The next step involves judging whether the models are adequate and some measurements need to be made for this purpose. It shall be highlighted that simple models are generally preferable over complex models, in accordance with the principle of parsimony.

Step 2: Estimation of model

This step involves estimating the coefficients (parameters) of all tentative models by minimizing the sum of squares of the fitting errors. The maximum likelihood and non-linear least squares methods are used to identify the parameters which need to be estimated. The non-linear least squares method is simply an algorithm that determines the minimum of the sum of squared error function. Once the parameters and their standard errors are known, the t value can be determined for each parameter and interpreted. The parameters that are significantly different from zero are included in the fitted model. Conversely, parameters that are not significant are excluded from the model.

Computing the residual mean square error is also useful to assess the fit and compare different models. The residual mean square error is also used to calculate forecast error limits. The formula used to calculate the residual mean square error is given by:

$$S^2 = \frac{\sum_{t=1}^n e_t^2}{n-r} = \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n-r} \quad (3.38)$$

where:

$e_t = Y_t - \hat{Y}_t$ = Residual at time t

n = Number of residuals

r = Total number of parameters estimated.

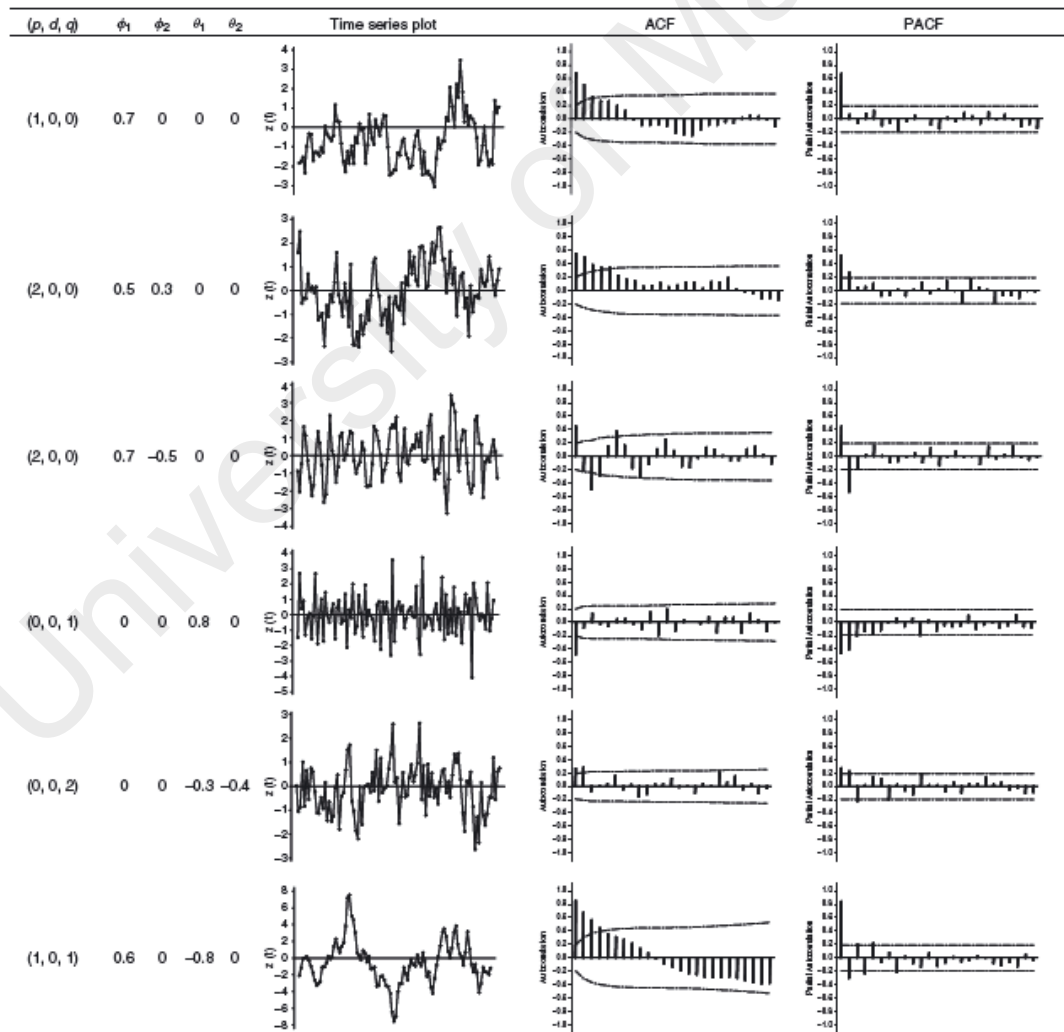


Figure 3.4: Examples of ACFs and PACFs of common ARMA models
Source: Bisgaard & Kulahci (2008)

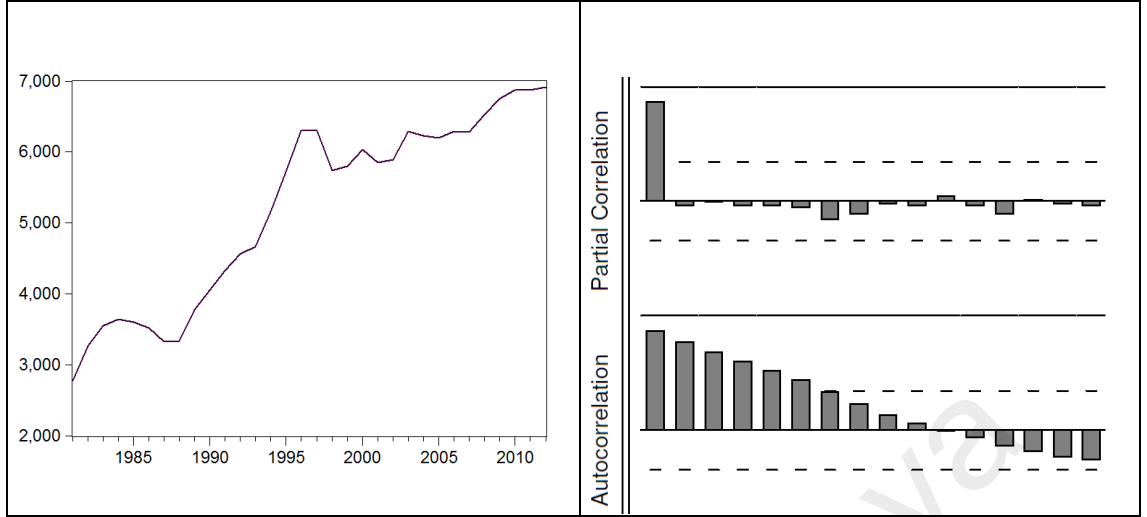


Figure 3.5: Example of a non-stationary data series and its ACF, PACF pattern

Step 3: Model checking

A model should be checked for its adequacy before it is used for forecasting. An adequate model is obtained when the residuals cannot be used to improve the forecasts. In other words, the residuals should be random. Histograms and normal probability plots are useful to check for normality whereas time sequence plots are useful to check for outliers. The criteria used to check the model are listed as follows:

- 1) The individual residual autocorrelations $r_k(e)$ should be small and generally within $\pm \frac{2}{\sqrt{n}}$ of zero.
- 2) The residual autocorrelations as a group should be consistent with those produced by random errors. Checks on adequacy can be done using chi-square (X^2) test based on the Ljung-Box Q test statistic. The residual autocorrelations as a group is examined using this test. The Q test statistic is given by:

$$Q = n(n+2) \sum_{k=1}^m \frac{r_k^2(e)}{n-k} \quad (3.39)$$

where:

n = Number of residuals

k = Time lag

m = Number of time lags included in the test

$r_k(e)$ = Sample autocorrelation function of the residuals lagged k time periods.

The Q test statistic is approximately distributed as a chi-square random variable with $m - r$ degrees of freedom, where r is the total number of parameters estimated in the ARIMA model. If the p value associated with Q is small (which is the case in this study where $p < 0.05$ and 0.1), the model is considered to be inadequate. Hence, a new model or a modified model needs to be developed until a satisfactory model is achieved.

Checks for normality can be done using several methods. The SAS software provides various test methods for this purpose, namely Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling. The Kolmogorov-Smirnov test for normality is used in this study. In this method, the normalized cumulative periodogram (J_k) is compared with the cumulative distribution function of a uniform(0,1) random variable. The normalized cumulative periodogram, F_j , of the series is given by:

$$F_j = \frac{\sum_{k=1}^j J_k}{\sum_{k=1}^m J_k}, \quad j = 1, 2, 3, \dots, m - 1 \quad (3.40)$$

where $m = n/2$ if n is even or $m = n-1/2$ if n is odd. The test statistic is the maximum absolute difference of the normalized cumulative periodogram and the uniform cumulative distribution function. For $m - 1$ greater than 100, if the Bartlett's Kolmogorov-Smirnov statistic exceeds the critical value $\alpha/(m - 1)^{1/2}$ (where $\alpha = 1.36$ or $\alpha = 1.63$ corresponding to 5% or 1% significance level, respectively), the null hypothesis that the series represents white noise will be rejected. The critical values for $m - 1 < 100$

can be found in the table of significance points of the Kolmogorov-Smirnov statistics with a sample size of $m - 1$ (SAS Institute Inc., 2004).

Step 4: Forecasting with the model

Forecasts are usually made using a satisfactory model or a model that has been determined to be the best model. Associating the confidence level to determine the upper and lower bounds is also useful to observe the variations that have occurred. However, the prediction interval may be an issue in this step. In general, the longer the forecast lead time, the longer the prediction interval. It is important to monitor the forecast errors in this step to ensure that the model provides reliable forecasts. If the magnitudes of the most recent errors tend to be consistently larger than previous errors, the model needs to be re-evaluated. The model also needs to be corrected if the recent forecast errors tend to be consistently positive (underprediction) or negative (overprediction).

In most cases, a model having the best criteria results in unstable forecasts. For this reason, modelling is first carried out using a model with only a few parameters. The need for additional parameters will be evident from an examination of the residual autocorrelations and partial autocorrelations. An MA parameter shall be added if the MA behaviour is apparent in the residual autocorrelations and partial autocorrelations. Conversely, the AR parameter will be increased if the residual autocorrelations tend to show an AR process. Moreover, the least squares estimates of ARIMA models tend to be highly correlated. When there are more parameters than necessary, this leads to 'trade-offs' among the parameters as well as unstable models that will produce poor forecasts.

3.11.1.5 Model selection criteria

It is important to identify the ARIMA model by examining a plot of the series and matching their sample autocorrelations and partial autocorrelation patterns. However, further analysis is required to select a satisfactory model. This can be done by computing the mean square error for each tentative model. A model which generates the smallest mean square error is generally preferable. In this study, the AIC and SBC are used to select the best model, and are computed using Equations (3.41) and (3.42). The best model is the one which has minimum AIC and SBC.

$$AIC = \ln \hat{\sigma}^2 + \frac{2}{n} r \quad (3.41)$$

$$SBC = \ln \hat{\sigma}^2 + \frac{\ln n}{n} r \quad (3.42)$$

where:

\ln = Natural log

$\hat{\sigma}^2$ = Residual sum of squares divided by the number of observations

n = Number of observations (residuals)

r = Total number of parameters (including the constant term) in the model.

The second term of Equations (3.41) and (3.42) is termed the penalty factor which is used to determine the inclusion of additional parameters in the model. In most cases, the AIC and SBC will produce the same results. However, since the SBC criterion imposes a greater penalty for a number of parameters compared to the AIC criterion, selecting a model with minimum SBC will result in a simpler model.

3.11.1.6 Summary of the ARIMA model construction

The availability of various statistical software in recent years has greatly facilitated time series modelling and reduced processing time. This is especially the case in determining the order of the ARIMA model, i.e. $AR(p)$ and $MA(q)$. The steps involved in the development of the ARIMA model are summarized as follows.

- Step 1. Obtain the time series plot of the data series in order to determine whether the series is stationary.
- Step 2. Check the lambda (λ) value of the Box-Cox transformation in order to determine whether transformation is required to attain a data series that is stationary in variance.
- Step 3. Conduct the ADF test to ensure that the series is also stationary in mean. Use the differencing technique if the series is not stationary even after the Box-Cox transformation has been employed. Obtain the order of differencing.
- Step 4. Develop the ARIMA model using statistical software once the data series is stationary.
- Step 5. Identify the order p and q . Use various combinations of p and q from 0 to 12. In this study, this step results in 169 different models.
- Step 6. Screen all of the models produced in order to obtain models that have a p -value less than 0.1 for all parameters. The intercept may be removed in order to ensure that all parameters of the model have a p -value less than 0.1.
- Step 7. Sort the competing models based on their SBC value, from the lowest to the highest.
- Step 8. Select the model with the lowest SBC as the tentative model until it fulfils the adequacy checks.

Step 9. Perform adequacy checks for the tentative model by evaluating the Q test statistic. Remove the model if the model does not fulfil this check, and continue checking the model with the second lowest SBC and so forth until the adequate (best) model attained.

Step 10. Use the best model for forecasting.

3.12 ARMA(p, q) intervention analysis

This technique is used to evaluate the effect of external events on a time series. In general, there are two patterns of intervention: step function and pulse function. The step function represents an intervention occurring at time T that remains in effect thereafter. In contrast, pulse function is an intervention that takes place at only one time period (Wei, 2006). The step function is given by:

$$S_t^{(T)} = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases} \quad (3.43)$$

while the pulse function is given by

$$SP_t^{(T)} = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases} \quad (3.44)$$

According to Enders (2003), there are several important extensions to the intervention. A more general ARMA (p, q) intervention model has the following mathematical representation:

$$Y_t = a_0 + A(L)Y_{t-1} + c_0 z_t + B(L)\varepsilon_t \quad (3.45)$$

where:

| | | |
|-------------------|---|---|
| Y_t | = | Response (dependent) variable at time t |
| Y_{t-1} | = | Response variable at time lags $t - 1$ |
| a_0, c_0 | = | Coefficients to be estimated |
| z_t | = | Intervention (or dummy) variable that has a value of zero prior to time t , $(t - 1)$, and unity beginning from time t |
| ε_t | = | White-noise disturbance |
| $A(L)$ and $B(L)$ | = | Polynomials in the lag operator L . |

There are several possible ways in which an intervention function is modelled which are listed as follows.

- (1) *Pure jump function:* The value of the intervention sequence jumps from 0 to 1 (unity), as shown in Figure 3.6(a).
- (2) *Impulse function:* The function z_t is zero for all periods except for one particular period in which z_t is unity, as shown in Figure 3.6(b). This pulse function best characterizes a purely temporary intervention. The effects of the single impulse may last many periods due to the autoregressive nature of the Y_t series.
- (3) *Gradually changing function:* An intervention may not reach its full force immediately, and it will take several periods to complete an intervention. This type of intervention function is shown in Figure 3.6(c).
- (4) *Prolonged impulse function:* Rather than a single pulse, the intervention may remain in place for one or more periods and then begin to decay. Since the road safety measures are allowed to reduce fatalities, the z_t sequence for the number of fatalities may be represented by a decaying function, as shown in Figure 3.6(d).

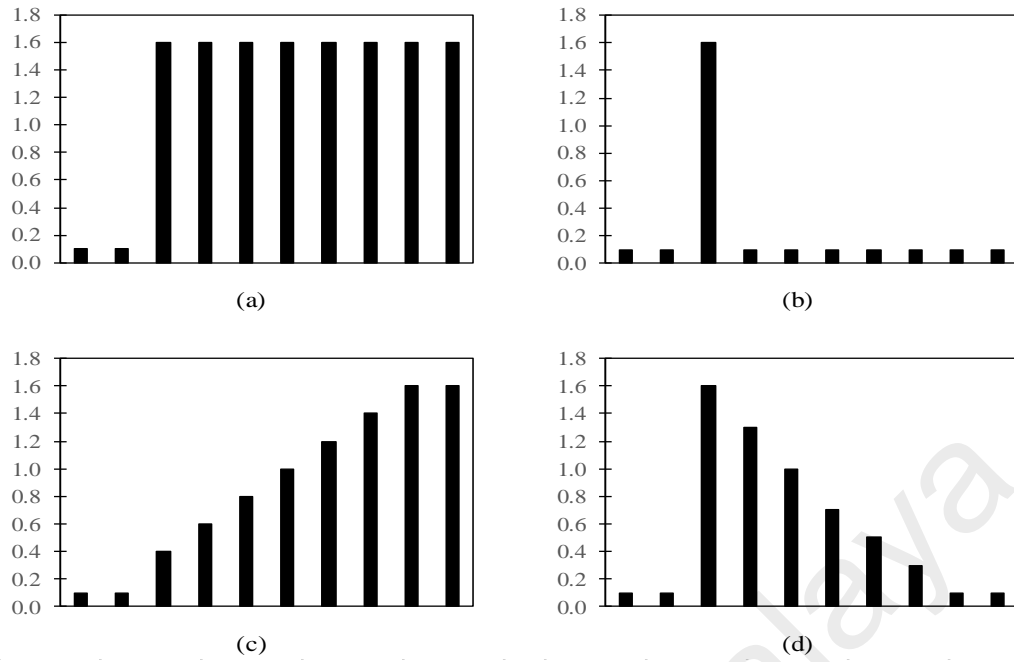


Figure 3.6: Typical intervention functions

According to Enders (2003), if two ARIMA models (pre-intervention and post-intervention) appear to be different, it is likely that the autoregressive and moving average coefficients have changed. In most cases, there are insufficient pre-intervention and post-intervention observations to estimate two separate models and therefore, the best-fitting ARIMA model over the longest data span is used in this study.

The impact of road traffic safety measures often results in gradual changes of road traffic casualties. Hence, it can be expected that the trend of the casualties post-intervention follows a step function. If the intervention results in a decaying trend, the response to the step input will be as shown in Figure 3.7.

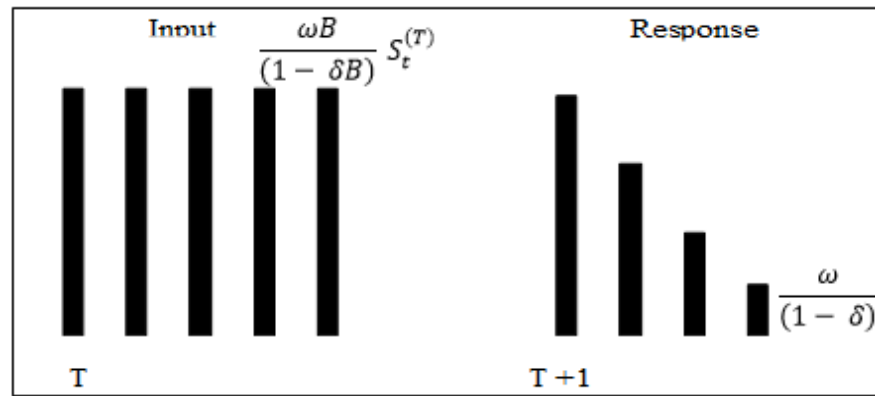


Figure 3.7: Typical response to a step input

3.12 Transfer function-noise model

It is typically assumed that effect of the intervention model is gained from a deterministic dummy variable. However, the effects of the safety measures (interventions) may be gained from something other than a deterministic dummy variable due to the fact that the factors which influence the level of the intervention may be different from time to time. For instance, the effect of the safety helmet law on the number and rate of road traffic casualties is influenced by the percentage of helmet usage. The level of enforcement also contributes significantly to the effectiveness of the law. Hence, it is assumed that the intervention variable can be any exogenous stochastic process. With this assumption, the transfer function-noise model is suitable to estimate the impact of the safety measures. In addition, the impact of legislations, standards, and road safety programmes follows a gradual, permanent step function of the intervention model. Consequently, the patterns follow the transfer function-noise model (Box & Jenkins, 1976), otherwise known as the dynamic regression model (Pankratz, 1991). The general representation of the transfer function-noise model is given by:

$$Y_t = \mu + \sum_i \frac{\omega_i(B)}{\delta_i(B)} B^{k_i} I_{i,t} + \frac{(1 - \theta_q B)}{(1 - \phi_p B)(1 - B)^d} a_t \quad (3.46)$$

where:

μ = Constant mean

$I_{i,t}$ = i^{th} input time series or a difference of the i^{th} input series at time t

k_i = Pure time delay for the effect of the i^{th} input series

$\omega_i(B)$ = Numerator polynomial of the transfer function for the i^{th} input series

$\delta_i(B)$ = Denominator polynomial of the transfer function for the i^{th} input series

a_t = white noise term.

The final part is the noise process that can be represented by the ARIMA model. The initial impact is given by ω_0 while the long-term effect (steady-state gain) is given by

$$\frac{(\omega_0 - \omega_1 B - \dots - \omega_i B^i)}{(1 - \delta_1 B - \dots - \delta_i B^i)}.$$

In general, the transfer function can be written as (Yaffee & McGee, 2000):

$$v(B) = \frac{(\omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s)}{(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)} I_{t-b} \quad (3.47)$$

The order of the transfer function refers to the levels of b , s and r . The order of b presents the delay or dead time, which is the time delay between incidences of changes in input, I_t , and the apparent impact on response, Y_t . The delay time b determines the pause before the input begins to have an effect on the response variable. The order of the regression is also represented by the values of s , which denotes the number of lags for unpatterned spikes in the transfer function. The order of decay is also denoted by the value of r . This parameter represents the patterned changes in the slope of the function. The order of this parameter signifies the number of lags of autocorrelation in the transfer

function. The denominator of the transfer function ratio consists of decay weights, δ_r , in which the time, t , varies from 1 to r . The magnitude of these weights controls the rate of attenuation in the slope. If there is more than one decay rate, the rate of attenuation may fluctuate.

There are several fundamental modelling strategies used in intervention analysis. In this study, the impact of the intervention is modelled on the entire data series before the residual noise is modelled. The SAS software is used to execute the transfer function-noise model, which involves determining the orders of the transfer function (b, s, r) and the noise process. The steps taken for intervention modelling are listed as follows (Box et al., 1994; Brocklebank & Dickey, 2003 and Montgomery et al., 2008).

- Step 1. Prewhiten the I_t series to attain stationary series, e.g., by some combination of Box-Cox and differencing transformation and generate the ARMA model. This step is similar to the step involved in ARIMA modelling.
- Step 2. Obtain the estimates of ω_b and δ_t .
- Step 2. Determine the orders of the transfer function, b, s and r . Tentatively use the sample cross-correlations function (CCF) or refer to the basic transfer function model structures shown in Figure 3.8. Box et al. (1994) have provided the details on calculating the CCF and the way to determine the orders of b, s and r based on the CCF. In this study, the CCF diagram from the software and the basic transfer function model structures are used to determine the orders.
- Step 4. Model the noise in order to determine ARMA(p, q).
- Step 5. Fit the overall model to attain the representation given in Equation (3.46).
- Step 6. Use SBC to select the best model among the competing models.
- Step 7. Check the adequacy of the model with the lowest SBC. If it is adequate, the model is selected as the best model. If the result reveals otherwise, repeat the

process using the model with the second lowest SBC and so forth until a model that fulfils the adequacy check is attained.

Step 8. Use the best model for forecasting.

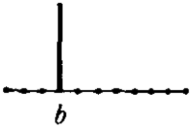
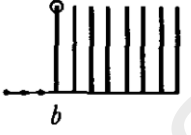
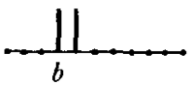
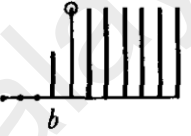
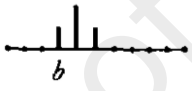
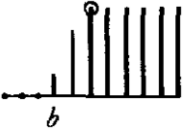
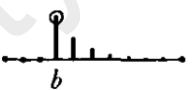
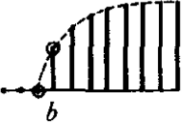
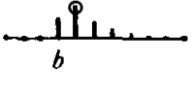
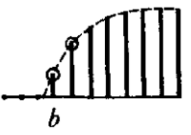
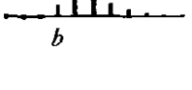
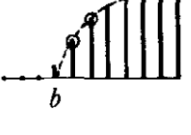
| r, s, b | Impulse response | Step response |
|-----------|---|---|
| 003 |  |  |
| 013 |  |  |
| 023 |  |  |
| 103 |  |  |
| 113 |  |  |
| 123 |  |  |

Figure 3.8: Examples of impulse and step response functions
Source: Adapted from Box et al. (1994)

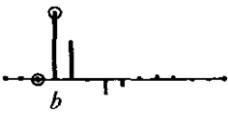
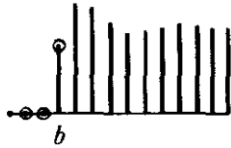

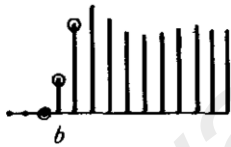

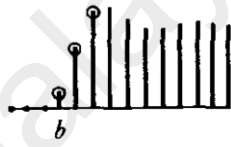
| r, s, b | Impulse response | Step response |
|-----------|---|---|
| 203 |  |  |
| 213 |  |  |
| 223 |  |  |

Figure 3.8, continued

3.13 State-space model

The state-space model is based on the Markov property, which implies the independence of the future of a process from its past, given the state of the present system. In a state-space model, the state of the process at a present time contains all of the past information that is required to predict the process behaviour in the future. According to Yaffee and McGee (2000), state-space model is a model that consists of stationary multivariate time series processes with dynamic interactions. This model is constructed from two basic equations. The state transition equation consists of a state vector of auxiliary variables as a function of a transition matrix and an input matrix, whereas the measurement equation consists of a state vector which is canonically extracted from observable variables. These vector models are estimated using a recursive protocol and they can be used for multivariate forecasting. Brockwell and Davis (2002) even

highlighted that state-space models are an extremely rich class of models for time series, which are well beyond the linear ARIMA and classical decomposition models.

A state-space model for a multivariate time series (Y_t) consists of two equations. The first equation is given by (Brockwell & Davis, 2002):

$$Y_t = G_t X_t + W_t \quad t = 1, 2, 3, \dots \quad (3.48)$$

This equation is known as the *observation equation* which expresses the w -dimensional observation Y_t as a linear function of a v -dimensional state variable X_t (the explanatory variables that have strong relationship to number of fatalities) and noise. $\{W_t\} \sim \text{WN}(0, \{R_t\})$ and $\{G_t\}$ is a sequence of $w \times v$ matrices. The second equation is known as the *state equation* or *state vector* that determines the state X_{t+1} at time $t + 1$ in terms of the previous state X_t and a noise term. The state vector is given by:

$$X_{t+1} = F_t X_t + V_t \quad t = 1, 2, 3, \dots \quad (3.49)$$

where $\{F_t\}$ is a sequence of $v \times v$ matrices. It shall be noted that $\{V_t\} \sim \text{WN}(0, \{Q_t\})$, and $\{V_t\}$ is uncorrelated with $\{W_t\}$ (i.e. $E(W_t V_s') = 0$ for all s and t). It is assumed that the initial state X_1 is uncorrelated with all of the noise terms $\{V_t\}$ and $\{W_t\}$.

The SAS software is used to determine the state vector and the state-space model representation. This procedure is similar to the modelling strategy proposed by Akaike (1976), which involves a canonical correlation analysis to identify the state-space model. Firstly, a sequence of unrestricted vector autoregressive (VAR) models are fitted and the AIC for each model is computed. The vector autoregressive models are estimated using the sample autocovariance matrices and *Yule-Walker* equations. The order of the VAR

model which yields the smallest AIC is chosen as the order (number of lags into the past) for use in the canonical correlation analysis.

The elements of the state vector are then determined from a sequence of canonical correlation analyses of the sample autocovariance matrices using the selected order. The sample canonical correlations of the past are computed with an increasing number of steps into the future. The variables that yield significant correlations are added to the state vector whereas those that yield insignificant correlations are excluded. Once the state vector is determined, the state-space model is fitted to the data. The free parameters in the F , G , and W matrices are estimated using the approximate maximum likelihood.

In summary, the following steps are taken to develop the state-space model.

- Step 1. Identify the independent and explanatory variables.
- Step 2. Perform stepwise regression analysis to obtain the explanatory variables that can be used to represent the model without collinearity problems.
- Step 3. Prewhiten the data series to attain stationary series, e.g., by some combination of Box-Cox and differencing transformation.
- Step 4. Execute the STATESPACE procedure in the SAS program to determine the state vector and the state-space model representations.
- Step 5. Select the state-space model representation and preliminary estimates.
- Step 6. Estimate the final state-space model representation.
- Step 7. Use the final model for forecasting.

In state-space modelling, the data series should be stationary. Checks for stationarity are conducted using the Augmented Dickey-Fuller (ADF) test. The differencing technique is then used to attain a stationary data series.

3.14 Chapter summary

ARIMA, transfer function-noise and state-space models will be used in this study. The multiple regression is used to select the explanatory variables that are correlated significantly with the number of road traffic fatalities. Then, these variables are the input variables for the state-space model. The effectiveness of a road safety measure is checked with the ARIMA and transfer function-noise model. Whilst, for forecasting the fatalities up to year 2020, the ARIMA, transfer function-noise and state-space models are employed, respectively.

The software packages are used to analyse the data, namely Microsoft Excel (2010), Minitab (Version 16), Statgraphics and SAS (Version 9). Minitab, Statgraphics and SAS are leading software packages in statistical analysis. Minitab is convenient for univariate analysis whereas SAS package is capable of performing numerous statistical operations. In addition, Eviews (Version 6) is used to crosscheck the unit root test results and is more likely to produce consistent results.

The forecasting period is chosen to be 8 years, which can be classified as medium-term forecast. In this study, it is deemed necessary to develop and evaluate several different models in order to determine the best forecasting model. The forecasting error serves as the basis for model construction and evaluation. Extrapolation is conducted to generate data from the observed historical values in order to check the accuracy of models and determine forecasting errors. Forecasting errors consist of the sum of the absolute values of errors, average forecast error and the sum of squared errors, which are used to compare the accuracy of a particular method from the alternative forecasting methods considered. These error patterns will reveal whether the method is suitable for forecasting, which will help an analyst to select the appropriate forecasting method.

In order to make a series stationary, either one of the following techniques can be used such as log transformation, power transformation, or Box Cox transformation techniques. Box-Cox transformation is the most commonly used variance-stabilizing transformation technique. When there is high autocorrelation in the data, the model itself changes rather than the levels, which often eliminates serial correlation. In this study, differencing is considered when the probability obtained from the ADF test is greater or equal to 0.1. Further analysis is required to select a satisfactory model. This can be done by computing the mean square error for each tentative model. A model which generates the smallest mean square error is generally preferable. In this study, the AIC and SBC are used to select the best model. The best model is the one which has minimum AIC and SBC. In most cases, the AIC and SBC will produce the same results. However, since the SBC criterion imposes a greater penalty for a number of parameters compared to the AIC criterion, selecting a model with minimum SBC will result in a simpler model.

CHAPTER 4: RESULTS AND ANALYSIS

This chapter consists of the key findings and statistical analysis of this study, beginning with descriptive statistics on the road safety scenario in Malaysia, followed by a detailed presentation of the results of the three models developed in this study. A list of the road safety measures and interventions implemented in order to improve road safety in Malaysia is also given in this chapter.

4.1 Key statistics on the road traffic safety scenario in Malaysia

The number and rate of selected road traffic casualties from the early 1980s until 2012 are examined in order to obtain a clear perspective of the road traffic safety scenario in Malaysia over the years. Simple linear regression using the least squares method is also superimposed on the plots in order to examine the trend in the number and rate of selected road traffic casualties. Nine types of road traffic safety fatalities, i.e. (1) fatalities according to the type of road users, (2) fatal accidents in urban and rural areas, (3) fatalities according to gender, (4) fatalities according to age group, (5) fatal accidents according to road category, (6) fatal accidents according to the type of road segment, (7) fatal accidents according to the type of junction, (8) fatalities according to traffic system and (9) fatal accidents according to the time of day are selected, and the number and rate of fatalities are stratified according to the type of road users.

Time series data based on the type of road users are found to be useful in road safety research in order to identify which component contributes significantly to the total number of fatalities. The number of fatalities per road user also provides insight on the trend of fatalities for road users, and serves as the groundwork to formulate strategies and tasks in a road safety plan. In addition, the rate and number of fatalities based on the types of road and road traffic system are also assessed. It shall be highlighted that the rate and

number of fatal accidents or driver/rider fatalities are generally preferable to reduce biases in comparative analysis. For example, if a fully occupied bus is involved in a fatal accident, the number of fatalities is dramatically higher compared with a motorcycle, even though both vehicles are involved in the same accident.

However, prior to knowing the trends of nine types of road traffic fatalities, it is first necessary to know the rate of road traffic fatalities per 10,000 registered vehicles and rate of road traffic fatalities per 100,000 population and these data are shown in Figure 4.1. These two parameters are commonly used to determine the level of safety achievement in countries throughout the world. The rate of fatalities per vehicle-kilometre travelled is also presented – however, these data are only available for the last few years. It can be observed from Figure 4.1 that the rate of road traffic fatalities per 10,000 registered vehicles and the rate of road traffic fatalities per 100,000 population shows opposing trends even though these parameters usually show a similar trend. The linear regression lines indicate that the rate of fatalities per 100,000 population increases whereas the rate of fatalities per 10,000 registered vehicles decreases from 1981 to 2012.

A thorough analysis is required to explain this unusual trend, but suffice it to say at the moment that this scenario is attributed to fact that the number of fatalities and rate of vehicle ownership constantly increase over the years in Malaysia. In normal cases, a country with a high vehicle ownership rate has a low number of fatalities. This is why the rate of fatalities per 10,000 registered vehicles is relatively low in Malaysia but it is still high per 100,000 population. However, the rate of fatalities per 10,000 registered vehicles is used for Malaysia in this study since this ratio takes into account the effect of the motorization level. This is also consistent with the road safety indicator suggested by Al Haji (2005), who developed a new method to rank the road safety achievement in ASEAN countries. Moreover, road safety authorities have targeted a rate of fatalities less than 2.0

per 10,000 registered vehicles by year 2010, as stated in the road safety plan launched in 2006. However, it is found that the rate of fatalities per 10,000 registered vehicles is above 3.0 in 2012, and therefore, the safety plan is underachieved.

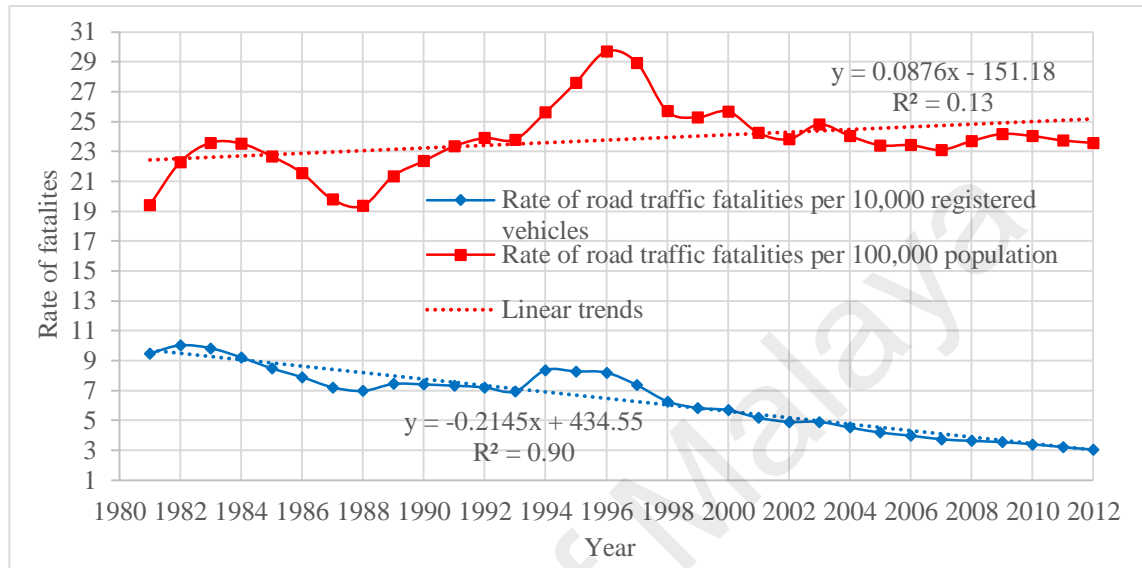


Figure 4.1: Rate of road traffic fatalities per 10,000 registered vehicles and rate of road traffic fatalities per 100,000 population due to road traffic accidents in Malaysia from 1981 to 2012

4.1.1 Road traffic fatalities according to the type of road users

The number of fatalities according to the type of road users is shown in Figure 4.2 whereas the rate of the fatalities of selected road users is shown in Figure 4.3. It can be seen that both the number and rate of fatalities due to road traffic accidents are the highest for motorcyclists. In 2012, the number of fatalities for motorcyclists accounts for 54% of the total number of fatalities, whereas the number of fatalities for motorcyclists and pillion riders account for 60% of the total number of fatalities. In the early years, the number of fatalities is second highest for pedestrians, but it is then replaced by motorcar occupants in recent years. It is evident from Figure 4.2 that the number of fatalities for most road users decreases over the years, with the exception of motorcyclists and

motorcar occupants. Indeed, the number and rate of fatalities for motorcar occupants have increased remarkably in recent years. The number of fatalities for motorcar occupants only contributes 8% of the total number of fatalities in 1983 but this figure has increased significantly to 21% in 2012. This dramatic increase occurs since 2007, which is unexpected since the rate of fatalities for motorcar occupants only shows a slight increase until year 2006. This may be one of the reasons why the safety target in 2010 is underachieved.

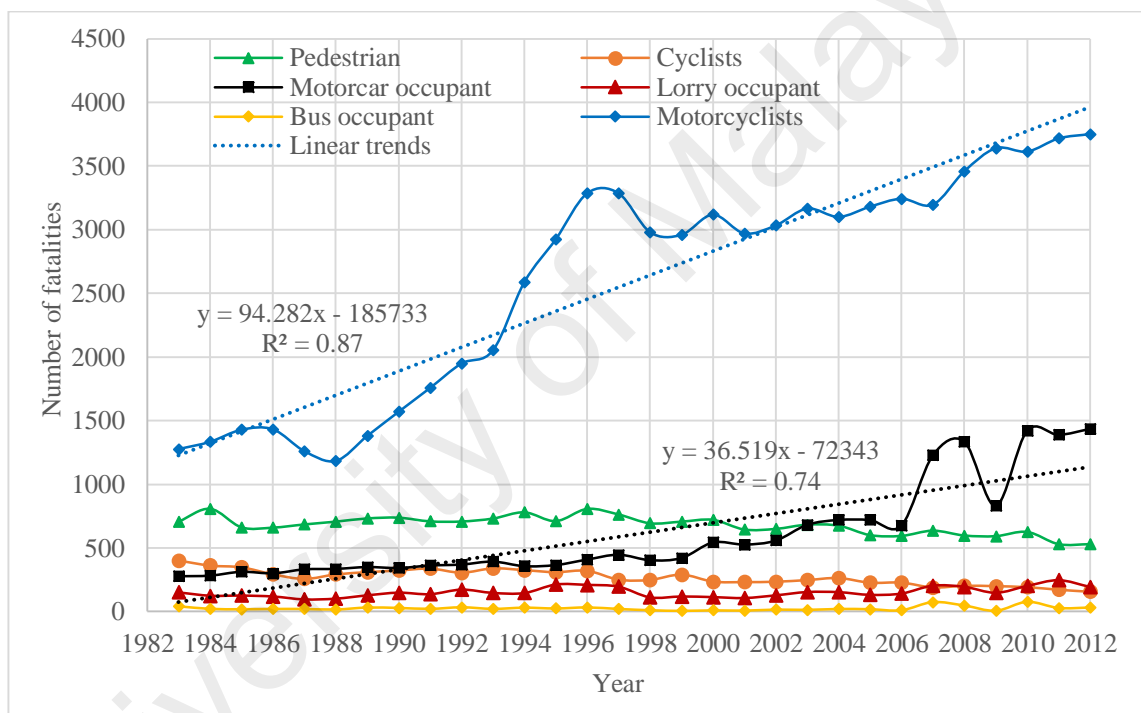


Figure 4.2: Number of fatalities according to the type of road users from 1983 to 2012

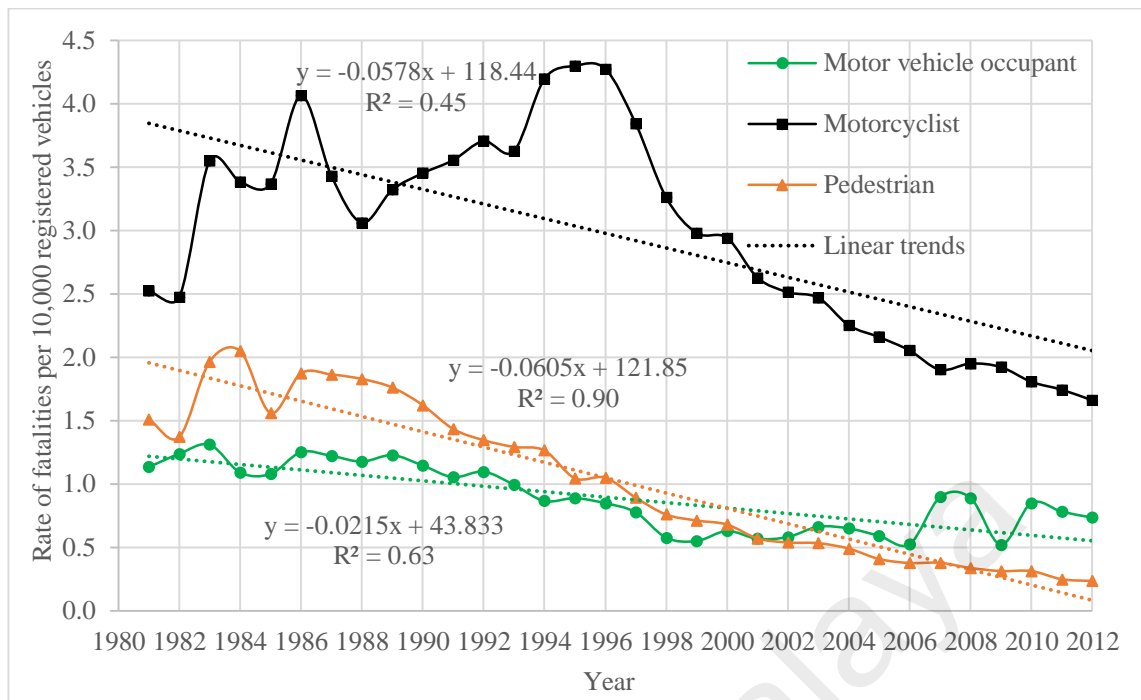


Figure 4.3: Rate of fatalities per 10,000 registered vehicles for selected road users from 1981 to 2012

4.1.2 Fatal accidents in urban and rural areas

Based on the population and housing census data published by the Department of Statistics (2011), the proportion of the population in urban areas has increased from 51 to 71% within 2000–2010. Thus, the number of population in urban areas is significantly higher than that in rural areas. The percentage of fatal accidents from 1993 to 2012 is shown in Figure 4.4 and ironically, it can be seen that the percentage of fatal accidents is higher in rural areas compared to urban areas. For instance, in 2012, the percentage of fatal accidents in rural areas is 69% whereas the remaining 31% is attributed to fatalities that occur on roads in urban areas.

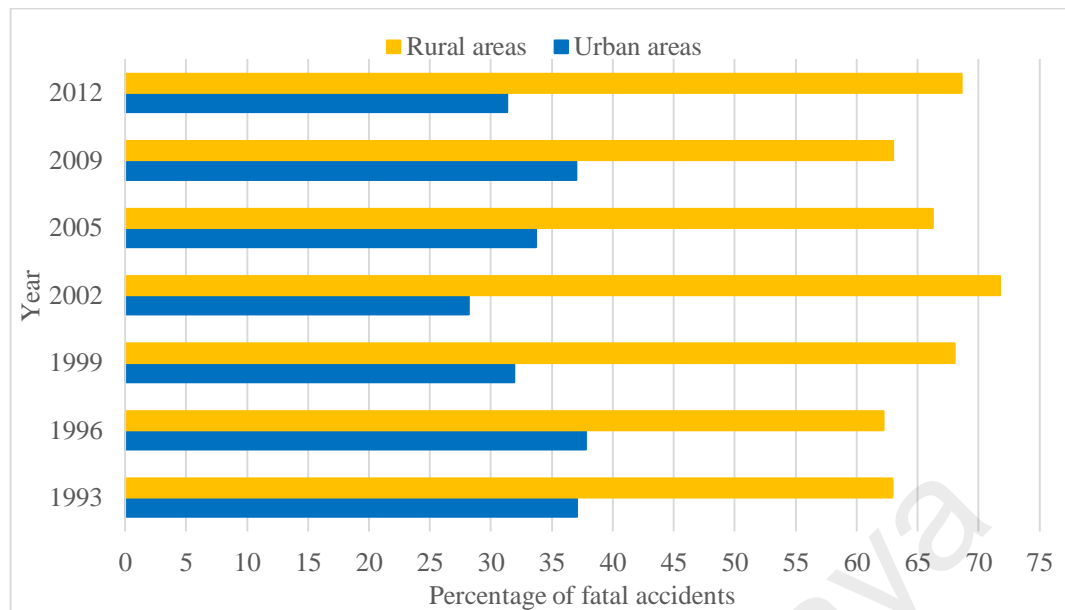


Figure 4.4: Percentage of fatal accidents from 1993 to 2012

Moreover, the trend in the rate of fatal accidents in urban and rural areas is consistent with the rate of fatal accidents per 10,000 registered vehicles. It can be seen that the rate of fatal accidents per 10,000 registered vehicles in rural areas is substantially higher than in urban areas, as shown in Figure 4.5. Nevertheless, even though the rate of fatal accidents declines in both urban and rural rate over the years, the rate of the accidents in rural areas shows a more pronounced decline. This indicates that there are more improvements with regards to road safety in rural areas compared to urban areas. A number of possible explanations regarding why the number of deaths is higher in rural areas compared to urban areas are given in Chapter 5.

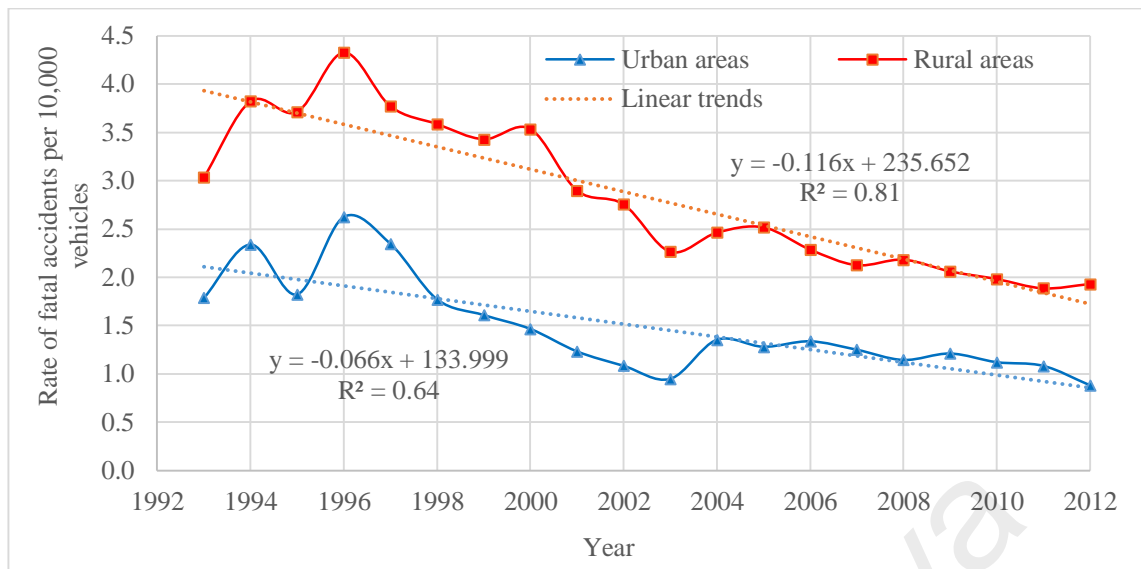


Figure 4.5: Rate of fatal accidents in urban and rural areas from 1993 to 2012

4.1.3 Driver/rider fatalities according to gender

The percentage of male and female driver/rider fatalities from 1993 to 2012 is shown in Figure 4.6. It can be seen that there is a stark contrast in the percentage of driver/rider fatalities between males and females, whereby there is an extremely high percentage of deaths of males due to road traffic accidents within this period. However, the proportion male and female drivers/riders is similar. In 2012, male drivers/riders constitute 48.5% of the total number of drivers/riders. Based on the road safety statistics, the number of male driver/rider fatalities accounts for more than 85% of the total number of fatalities. The rate driver/rider fatalities for males and females from 1993 to 2012 is shown in Figure 4.7, and it can be seen that there is a significant improvement in the rate of male driver/rider fatalities within 1993–2012, which is apparent from the declining trend. In contrast, the rate of female driver/rider fatalities is rather constant over the same period and is rather low compared to males.

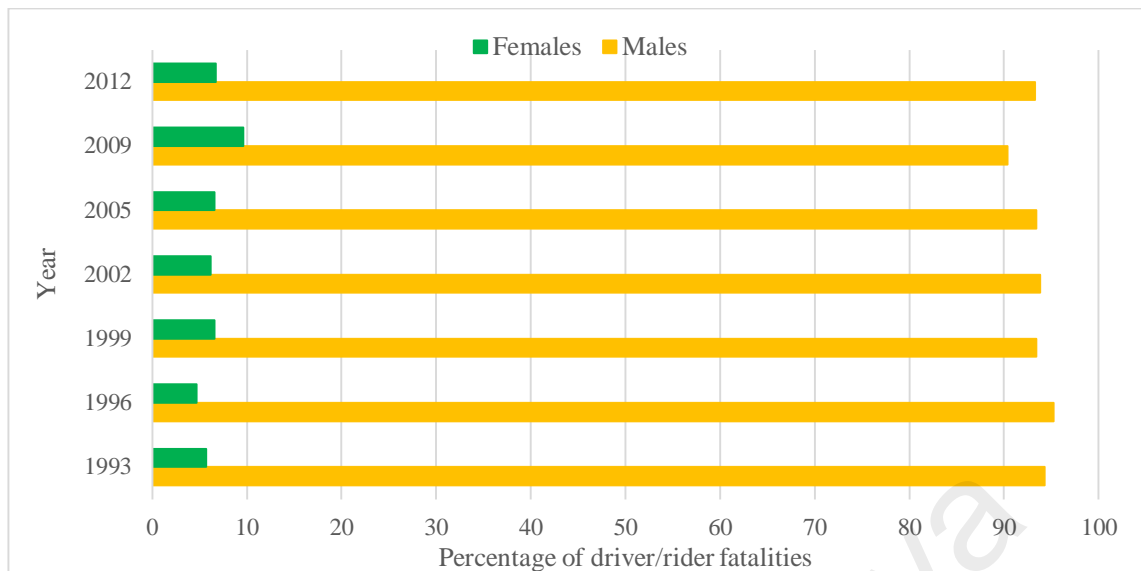


Figure 4.6: Percentage of male and female driver/rider fatalities from 1993 to 2012

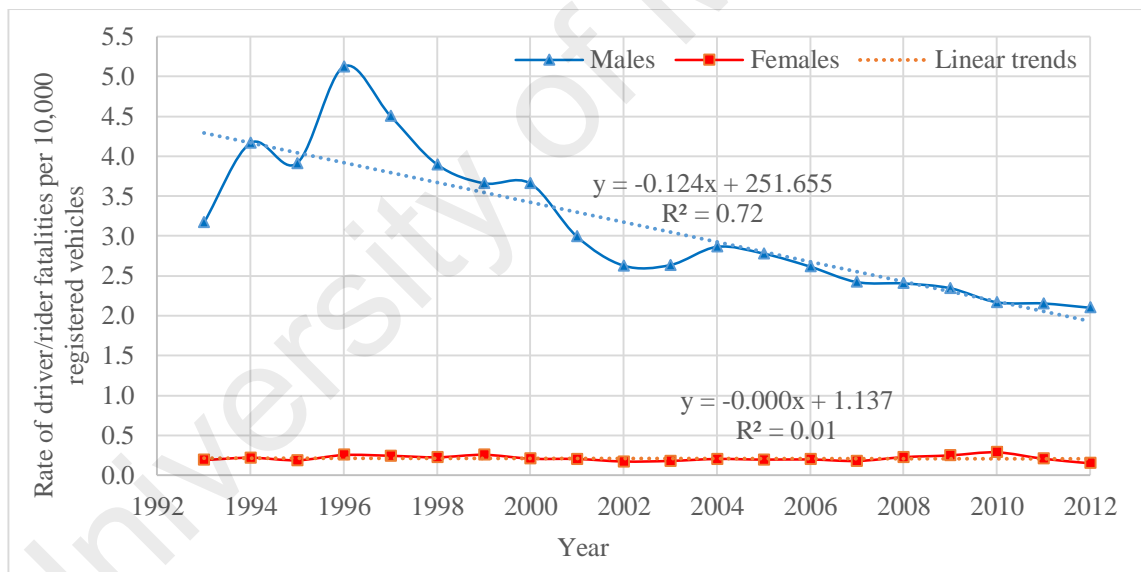


Figure 4.7: Rate of male and female driver/rider fatalities from 1993 to 2012

4.1.4 Driver/rider fatalities according to age group

The percentage of driver/rider fatalities according to age group from 1993 to 2012 is shown in Figure 4.8. It can be observed that drivers/riders aged 16–25 years constitute

the highest percentage of fatalities due to road traffic accidents. This age group accounts for 35% of the total driver/rider fatalities in year 2012, considering the fact that they only make up 17% of the total population. This is followed by drivers/riders aged 26–35 years (22%) and 36–45 years (13.5%), indicating that young adult drivers make up the highest number of victims involved in road traffic accidents. It can be observed that adult and older adult drivers/riders each constitute a relatively low percentage of fatalities. However, the percentage of fatalities involving drivers/riders aged 16–25 years seem to decline slightly over the years. The rate of driver/rider fatalities per 10,000 registered vehicles according to age group from 1993 to 2012 is shown in Figure 4.9. It is apparent that the rate of driver/rider fatalities decline significantly from 1996 to 2012 specifically for drivers/riders within the 16–25 years age group. The difference in the rate of driver/rider fatalities between this group and other age groups is relatively high from 1993 to 2000, but it is not as pronounced in recent years.

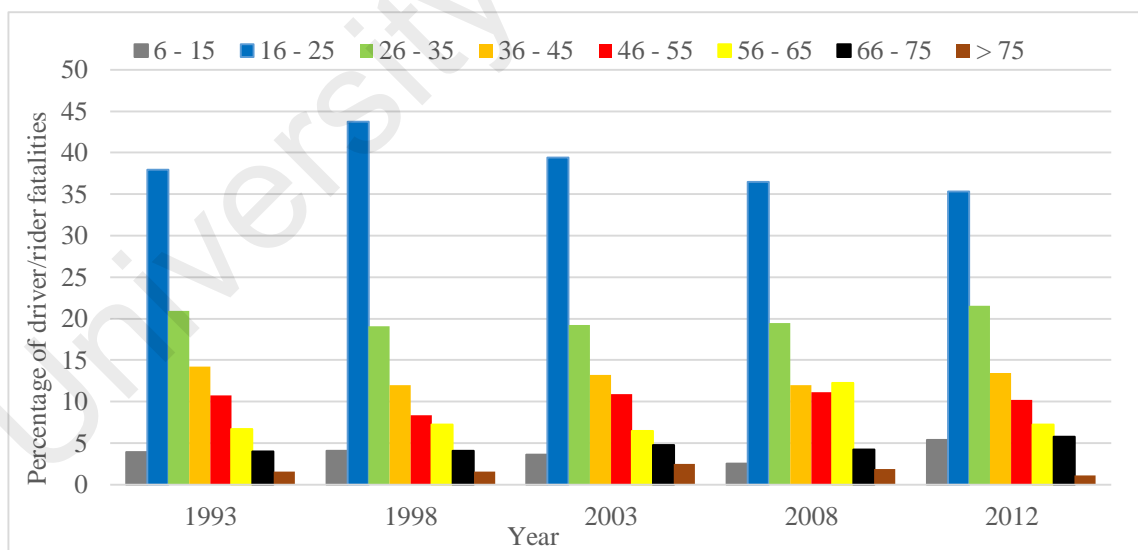


Figure 4.8: Percentage of driver/rider fatalities from 1993 to 2012 according to age group

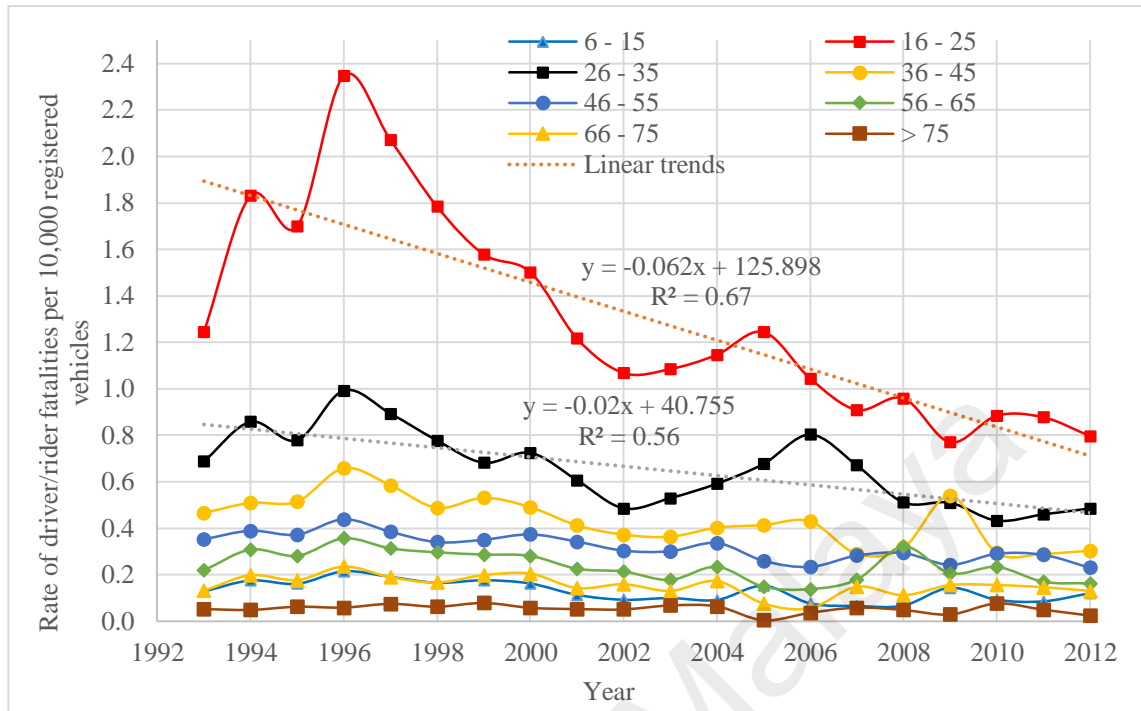


Figure 4.9: Rate of driver/rider fatalities from 1993 to 2012 according to age group

4.1.5 Fatal accidents according to road category

In 2012, the highest percentage of fatal accidents occur on federal roads (35%), considering the fact that federal road networks only constitute 14% of the total road length. This is followed by accidents on state roads, municipal roads and expressways, in which the number of driver/rider fatalities is 28, 17 and 11%, respectively. This trend occurs even in recent years, as shown in Figure 4.10. Hence, it is desirable to include the rate of fatal accidents per vehicle-kilometre travelled for comparison, but it is not possible to do so since the data for vehicle-kilometre travelled are unavailable. For this reason, the rate of fatal accidents per length of road (km) is used for comparison and the results are tabulated in Table 4.1. It is interesting to note that the highest rate of driver/rider fatalities per kilometre of road occurs on expressways, even though the highest percentage of fatal accidents in year 2012 occurs on federal roads (35%), as mentioned previously. Based on

the results shown in Figure 4.11, even though the rate of fatal accidents on various types of roads declines over the years, the rate of fatal accidents on expressways show neither a remarkable decrease nor increase. Rather, the rate of fatal accidents on expressways fluctuates within a range of 0.3–0.4 from 1993 to 2012.

Table 4.1: Rate of fatal accidents per kilometre of road in 2012

| Road category | Number of driver/ rider fatalities | Road length (km) | Rate of driver/rider fatalities per kilometre |
|-------------------------------|---------------------------------------|---------------------|---|
| Expressways | 704 | 1,742 | 0.404 |
| Federal roads | 2,252 | 17,474 | 0.129 |
| State and municipal roads* | 2,860 | 108,301 | 0.026 |

*: The Highway Planning Unit combines the length of both roads.

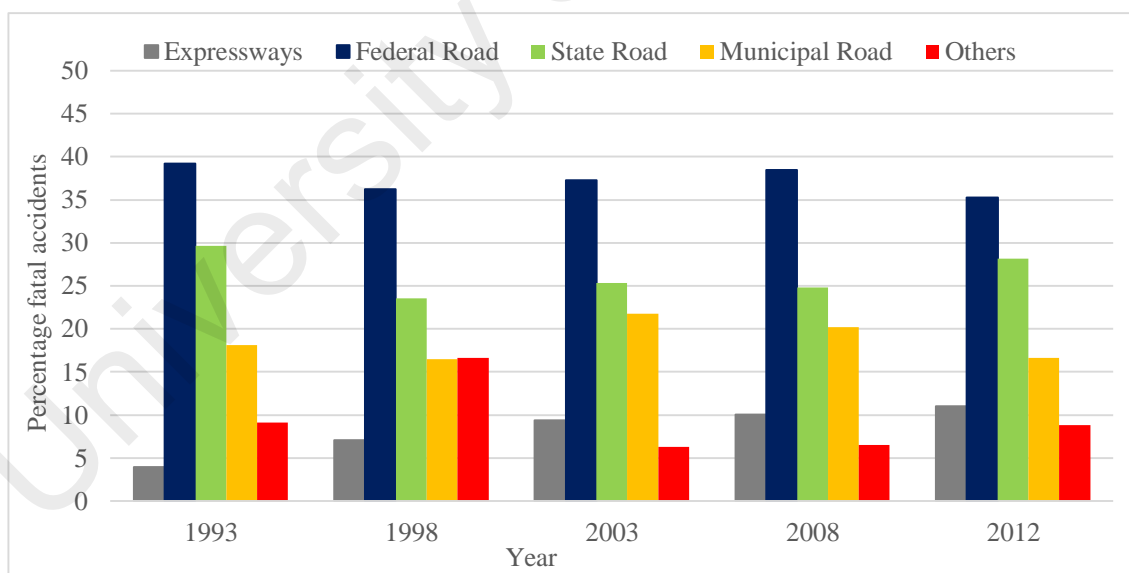


Figure 4.10: Percentage of fatal accidents from 1993 to 2012 according to road category

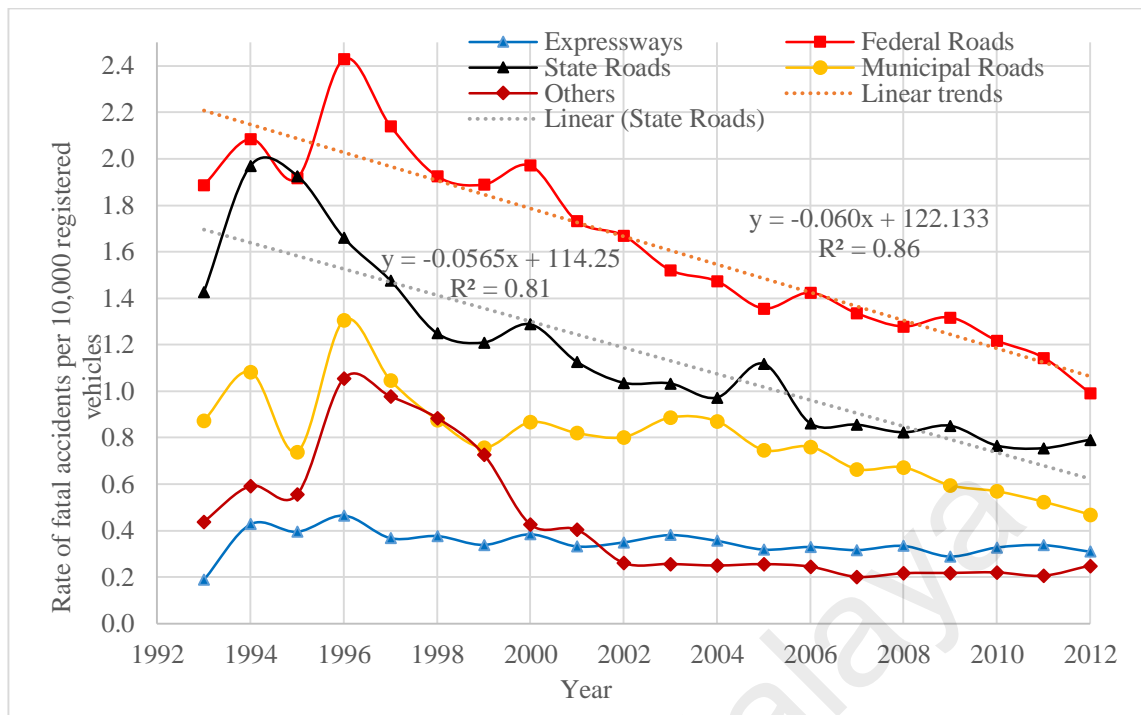


Figure 4.11: Rate of fatal accidents from 1993 to 2012 according to road category

4.1.6 Fatal accidents according to the type of road segment

In this study, it is found that the percentage of fatal accidents on straight road segments is considerably higher compared to that for bend segments regardless of the number of straight and bend road segments in Malaysia. It is observed that the percentage of fatal accidents that occur on straight segments increases from 1993 to 2012, as shown in Figure 4.12. It can be seen that there is a large percentage of fatal accidents that occur on straight road segments, with a value of 75–80%. However, according to Figure 4.13, the rate of fatal accidents decreases significantly from 1993 to 2012 for straight road segments compared with bend ones. This implies that there are major improvements in the safety features of straight road segments in Malaysia.

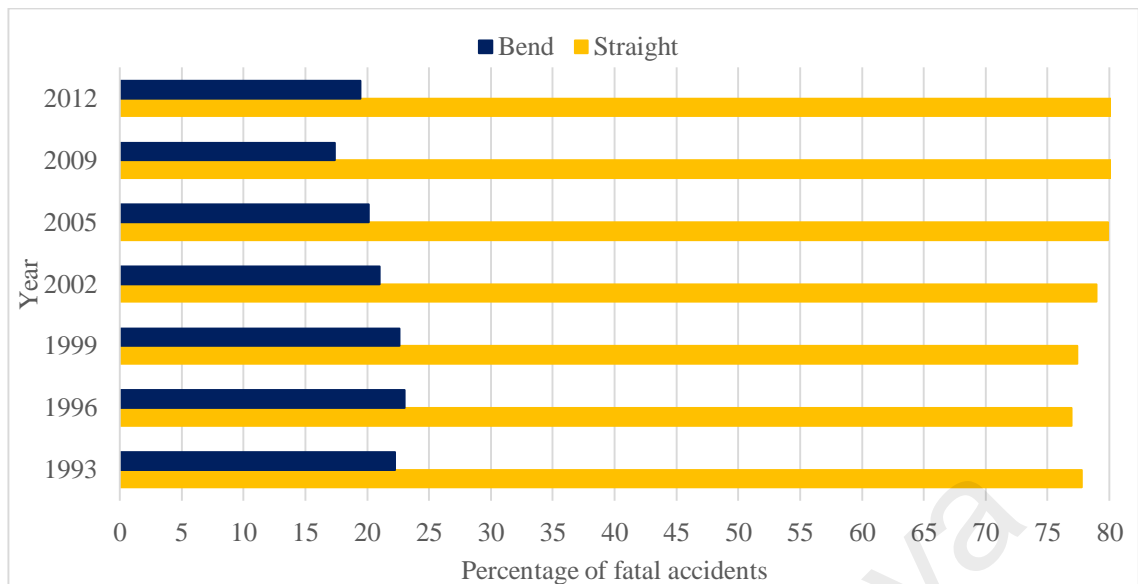


Figure 4.12: Percentage of fatal accidents from 1993 to 2012 according to the type of road segment

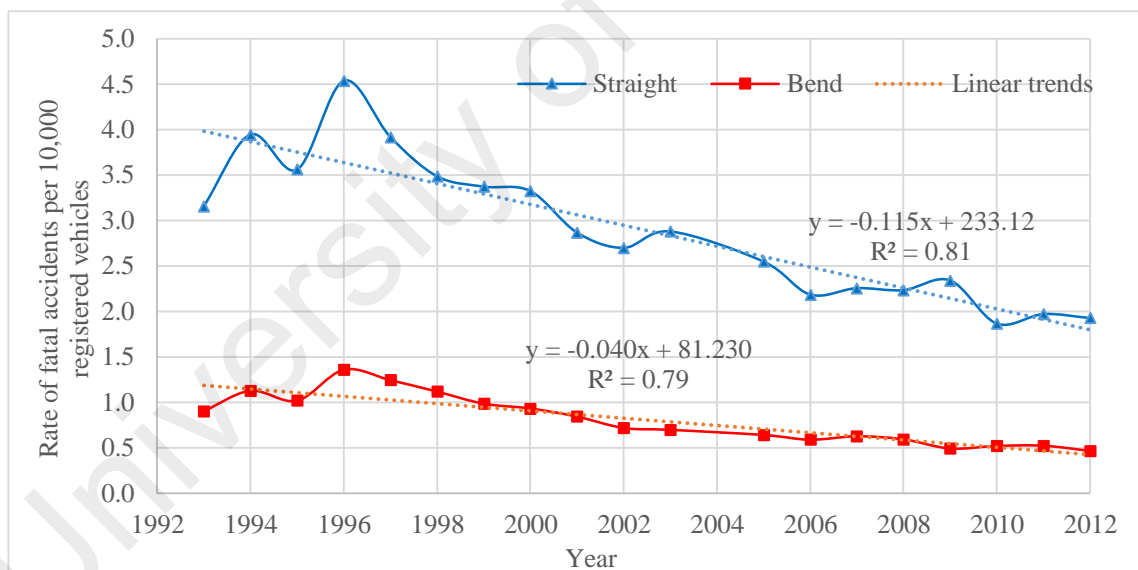


Figure 4.13: Rate of fatal accidents from 1993 to 2012 according to the type of road segment

4.1.7 Fatal accidents according to the type of junction

In this study, it is found that the highest percentage of fatal accidents occurs at T- and Y-junctions, regardless of the number of various types of junctions in Malaysia, as shown in Figure 4.14. This is followed by the percentage of fatal accidents at cross (four-legged) junctions. In contrast, the least percentage of fatal accidents occurs at interchanges and roundabouts. The high percentage of fatal accidents at T- and Y-junctions may be attributed to the skew angles of the junctions, which restricts the visibility of motorists. The rate of fatal accidents from 1993 to 2012 according to the type of junction is shown in Figure 4.15, and it can be observed that there is a substantial decrease in the rate of fatal accidents per 10,000 registered vehicles at T- and Y- junctions over the years. This leads to a small difference between the rate of fatalities at these junctions and other types of junctions in 2012. It is apparent that the rate of fatal accidents per 10,000 registered vehicles at interchanges and roundabouts is rather low, which may be due to the fact that these junctions have fewer conflict points.

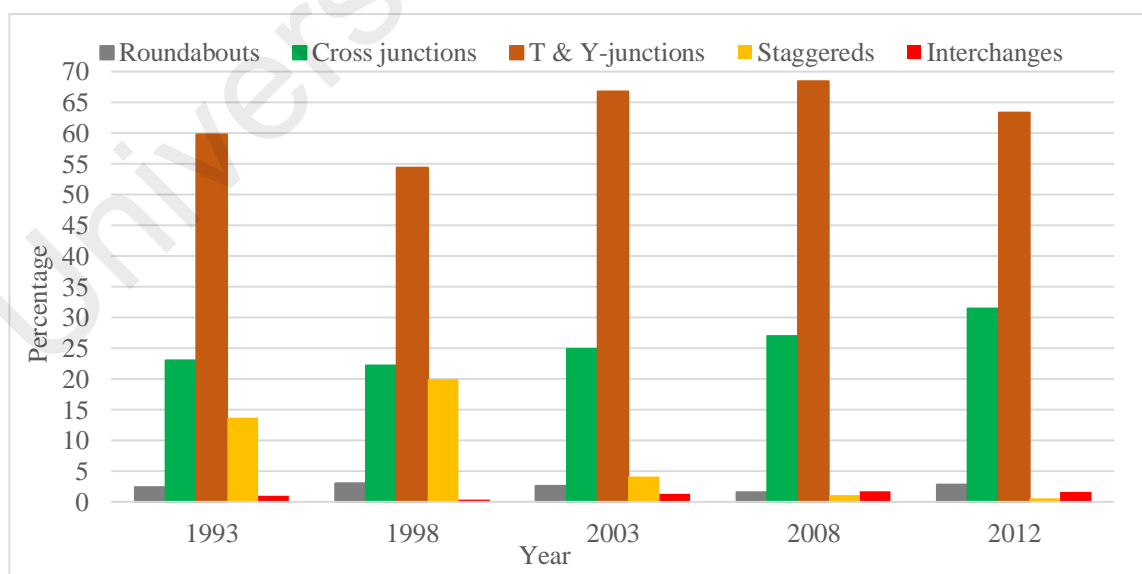


Figure 4.14: Percentage of fatal accidents from 1993 to 2012 according to the type of junction

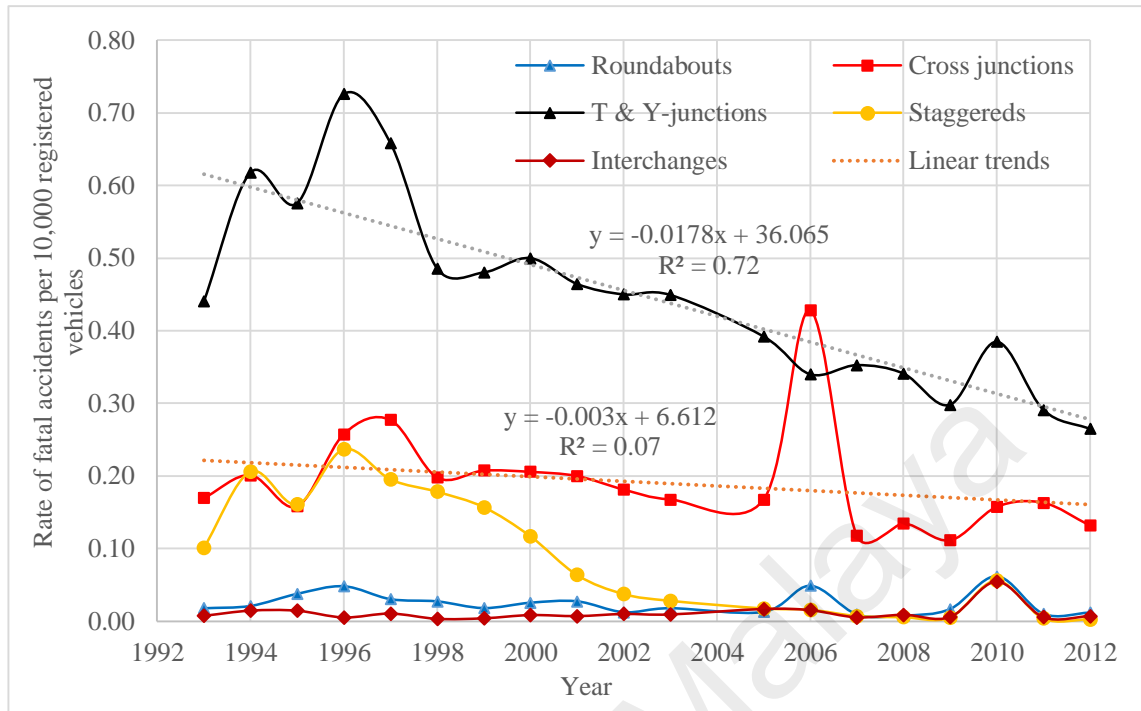


Figure 4.15: Rate of fatal accidents from 1993 to 2012 according to the type of junction

4.1.8 Driver/rider fatalities according to traffic system

Figures 4.16 and 4.17 highlight that single carriageway traffic system results in the highest percentage and rate of driver/rider fatalities. The percentage of driver/rider fatalities resulting from single carriageway roads accounts for more than 70% of the total number of fatalities. This is followed by the percentage of driver/rider fatalities for one-way roads. It is believed that the rate of driver/rider fatalities will decline in the future since increasing demands on road capacities have led the authorities to build more multi-lane dual carriageway roads. Upgrading works typically involve replacing existing single carriageway roads with dual ones. Figure 4.17 shows that there is a significant decline in the rate of fatalities per 10,000 registered vehicles for single carriageway roads from 1993 to 2012. There are two possible explanations for this. Firstly, the number of single carriageway roads has decreased because these roads are replaced with dual ones. It is

believed that segregating the roadways, specifically with appropriate median, will reduce the number of accidents. Secondly, the safety features of single carriageway roads are improved by the implementation of blackspot treatment programmes and upgrading works based on the recommendations of safety auditors.

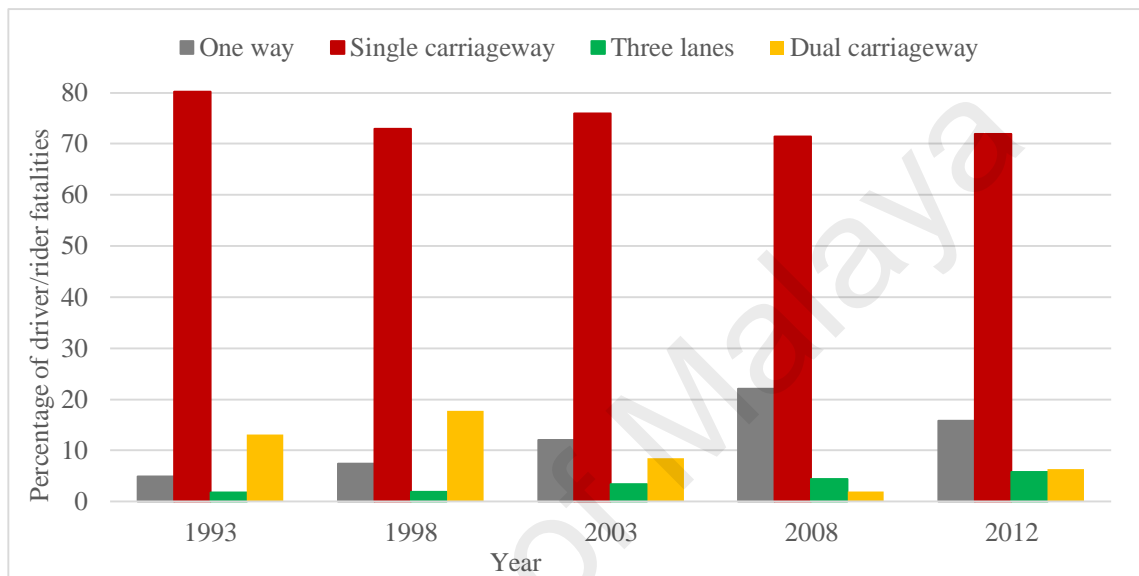


Figure 4.16: Percentage of driver/rider fatalities from 1993 to 2012 according to traffic system

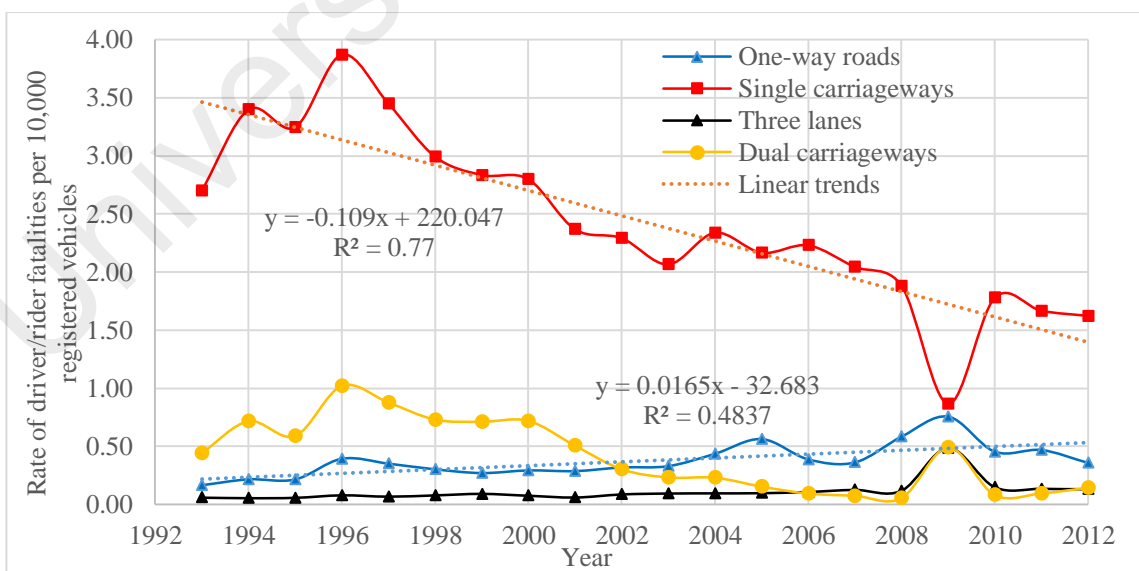


Figure 4.17: Rate of driver/rider fatalities from 1993 to 2012 according to traffic system

4.1.9 Fatal accidents according to the time of day

In 1993, the ratio of the percentage of fatal accidents between daytime and night time is 60% : 40%. It can be observed that the percentage of fatal accidents that occurred during daytime decreases from 1993 to 2012, as shown in Figure 4.18. In 2012, the ratio of the percentage of fatal accidents between daytime and night time is 53% : 47% – however, it is apparent that the percentage of fatal accidents during daytime is still predominant. It is evident from Figure 4.19 that the rate of fatal accidents per 10,000 registered vehicles decreases over the years for both daytime and night time, and the difference between these two rates is insignificant in year 2012.

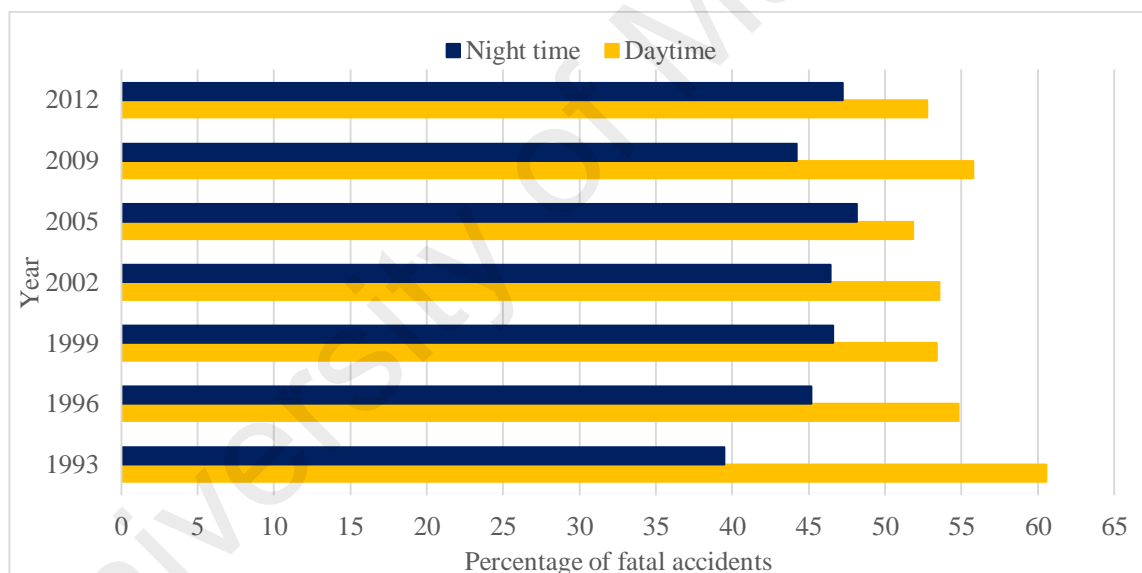


Figure 4.18: Percentage of fatal accidents from 1993 to 2012 according to the time of day

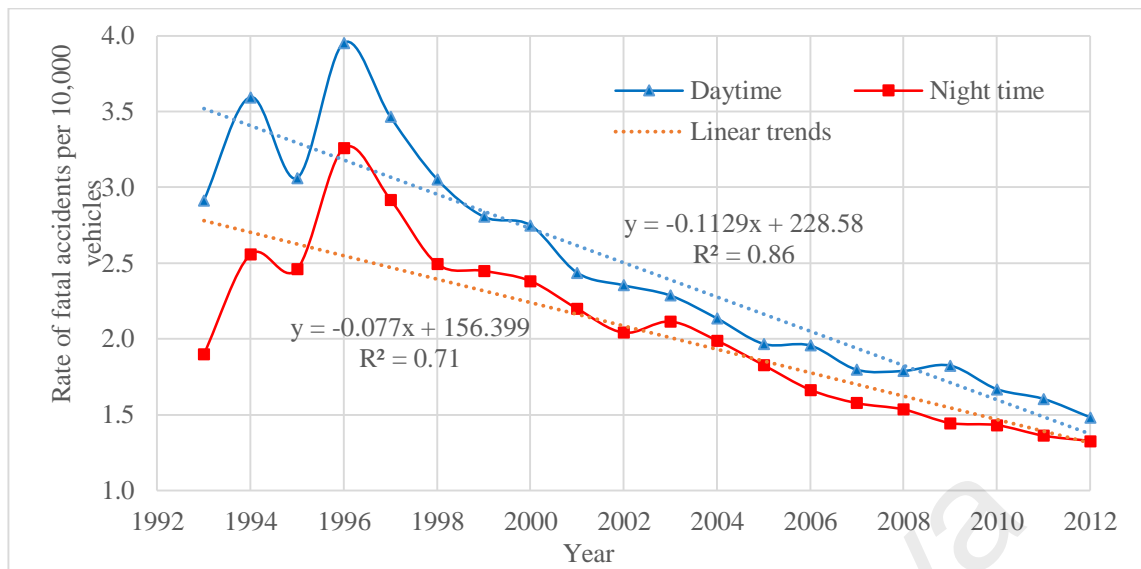


Figure 4.19: Rate of fatal accidents from 1993 to 2012 according to the time of day

4.2 Investigation of the government interventions and road safety measures on road safety improvement

In Malaysia, various interventions or safety measures have been implemented over the years to improve road safety. These interventions are implemented through legislations, standards, guidelines and safety programmes including those specified in road safety plans. The various interventions, including tasks which support the interventions of road safety, are listed as follows. The earliest interventions, however, are not included since the data and statistics on road traffic accidents are only available since the late 1960s.

4.2.1 Legislations on the road traffic safety

- (1) The Road Traffic Ordinance 1958 came into effect on July 1959.
- (2) The Motorcycles (Safety Helmet) Rules 1973 came into force on 1st March 1973.
- (3) The Motorcycles (Safety Helmet) (Amendment) Rules 2012 came into operation on 1st September 2012.

- (4) The Motor Vehicles (Seat Belt) Rules 1978 for front seats of motor vehicles came into force on 1st January 1979.
- (5) The Motor Vehicles (Safety Seat Belts) (Amendment) Rules 2008 for passengers in the back seats of motor vehicles came into force on 1st January 2009.
- (6) Road Transport Act 1987 involves driving license, seat belt and facilities for pedestrian.
- (7) The Motor Vehicles (Speed Limit) Rules 1989 came into force on 1st February 1989.
- (8) The Motor Vehicles (Speed Limit) (Amendment) Rules 1998 came into force on 15th March 1998.
- (9) The Motor Vehicles (Construction, Equipment and Use) (Speed Monitoring Device) Rules 1998 came into force on 15th March 1998 and is only applicable in Peninsular Malaysia.
- (10) The Road Transport Act 1987 for driving while under the influence of intoxicating liquor or drugs came into force on 1st January 1988.
- (11) The Road Transport (Amendment) Act A1065 1999 for driving while under the influence of intoxicating liquor or drugs came into force on 1st October 1999. The BAC limit is specified in this Act, i.e. 0.8 g/l.
- (12) The Motorcycle Daytime Running Headlight Regulation was made compulsory on September 1992.
- (13) The Motor Vehicles (Periodic Inspection, Equipment and Inspection Standard) Rules 1995 came into effect on 1st January 1995.

4.2.2 Standards

- (1) Specifications for protective helmets for vehicle users MS 1:1969.
- (2) Specifications for protective helmets for vehicle users MS 1:1996.
- (3) Specifications for protective helmets for vehicle users MS 1-1:2011.
- (4) Specification for seat belts for motor vehicles in 2003 (MS 1175:2003).
- (5) Installation of lighting and signalling devices for motor vehicles and their trailers in 2004 (MS ISO 303:2004).
- (6) Rethreaded pneumatic rubber tyres for passenger cars and commercial vehicles specification in 2008 (MS 1394:2008) and 2010 (MS 224:2005 Amd. 1:2010).

4.2.3 Guidelines

- (1) *Arahan Teknik Jalan* (ATJ) published by the Public Works Department (JKR), Ministry of Works since 1985.
- (2) Guidelines published by the Road Engineering Association of Malaysia (REAM) since 2002.
- (3) Guidelines for safety audit of roads and road projects in Malaysia published by the Public Works Department (JKR), Ministry of Works in 1997.

4.2.4 National road safety targets and plans

- (1) National Road Safety Target in 1991.
- (2) National Road Safety Target 1996–2000.
- (3) Road Safety Plan 2006–2010 launched on 15th May 2006.
- (4) Road Safety Plan of Malaysia 2014–2020.

4.2.5 Programmes

- (1) Development of exclusive lanes for motorcycles in the early 1970s and extensive development of such lanes since 1993.
- (2) Blackspots treatment programme since 1995.
- (3) Wide coverage media campaigns since 1997.
- (4) Road safety auditing for federal route projects since 1998 and state routes since 2007.
- (5) Integrated Road Safety Operation during festive seasons (*Ops Sikap*) by the Royal Malaysian Police (PDRM), the Road Transport Department (JPJ) and other government agencies since 2001.

4.2.6 Formation of government agencies

- (1) Road Safety Cabinet Committee established in 1989.
- (2) Road Safety Section in the Public Works Department (JKR), Ministry of Works, established in October 1997.
- (3) Road Safety Department (JKJR) in the Ministry of Transport, established on 15th September 2004.
- (4) Malaysian Institute of Road Safety Research (MIROS), established on 3rd January 2007.

The effectiveness of road safety interventions at the macro level can be measured by analysing the casualties before and after the enactment of legislations, national safety measures and standards. On the contrary, the impact of guidelines and programmes is typically at the micro level and it takes a considerable amount of time (several months or years) for these guidelines and programmes to be implemented throughout the nation.

Moreover, the effect of the national road safety plans may be relatively difficult to measure since they comprise a number of different programmes including legislations. Nevertheless, one of the objectives of this study is to investigate the impact of road safety plans owing to their significance. Furthermore, some of the legislations, as mentioned by Elvik et al. (2009), have effect at both the individual and aggregate levels. For instance, it can be expected that the Act which makes it mandatory for drivers to fasten their seat belts would reduce the probability of deaths for drivers (individual level). At the same time, it can be expected that this Act will increase the percentage of seat belt usage, which in turn reduces the number of fatalities in the country (aggregate level).

In this study, the Box-Jenkins (ARIMA) and the intervention method (the transfer function-noise model) are used to investigate the effectiveness of road safety measures (interventions) in Malaysia. The number and rate of casualties or accidents before and after the implementation of safety measures or legislations were made compulsory are analysed. Attempt has been made to use the longest span for the years of observation. Nevertheless, there are a number of conditions which limit the number of observations and span of the data series, and are listed as follows:

- (1) Most of the data regarding the number of accidents and casualties are published annually.
- (2) A number of safety measures and interventions were launched before 1981. However, statistics on national road safety are only available from year 1981 onwards. It shall be noted that before 1981, these statistics are available only for Peninsular Malaysia.
- (3) The effect of an intervention on the number of casualties observed is examined until the year before another major intervention (nationwide scope) is

implemented. This is done to prevent the overlapping of measures or interventions which will influence the interpretation of the casualties observed.

- (4) Various safety measures and interventions had been applied extensively in 1997 – the same year in which Malaysia is burdened with a regional economic crisis. For this reason, post-intervention analysis of certain observations is carried out until 1996.
- (5) Since 2001, integrated enforcement has been carried out by agencies during festive seasons. This year is also considered to be a time in which safety measures are affected by other safety measures.
- (6) A large number of observations have been carried out in recent years. For this reason, it is deemed impractical to examine observations within a relatively short time period.

Owing to the limitations discussed above, it is not possible to analyse the impact of all safety measures and interventions. In this study, a statistical software is used to determine whether a data series is adequate or not for analysis. The tentative interventions (road safety measures) included in the modelling are listed as follows:

- (1) Motorcycles (Safety Helmet) Rules in 1973
- (2) Motor Vehicles (Seat Belt) Rules in 1978
- (3) Road Transport Act 1987 for Drunken Driving in 1988
- (4) Motor Vehicles (Speed Limit) Rules in 1989
- (5) Motorcycle Daytime Running Headlight Regulation in 1992
- (6) Specifications for Protective Helmets for Vehicle Users (MS 1:1996)
- (7) Road Safety Programmes in 1997 following the Road Safety Target 1996–2000
- (8) Integrated Road Safety Operations (*Ops Sikap*) since 2001
- (9) Road Safety Plan 2006–2010

4.3 Modelling the impact of the interventions

The impact of an intervention on road traffic safety is measured using the transfer function-noise model. ARIMA model is also developed to forecast road traffic safety without such an intervention. The forecasted casualties obtained from the transfer function-noise model are compared with those obtained from the ARIMA model in order to assess the impact of the interventions. Both modelling techniques are classified under univariate time series analysis.

4.3.1 Data for the univariate time series analysis

The first step involved in the modelling process is to determine which the road traffic safety data are relevant to each safety measure. The number of road traffic fatalities is given the priority to relate the impact of safety measures implemented with road traffic safety. However, the number of accidents is used if the number of fatalities in relation to a safety measure is unavailable. Preliminary analysis on the safety data available reveals that only the impact of certain safety measures can be linked to the number of fatalities while others cannot. In such a case, the number of accidents is used in the analysis.

In addition, the rate of fatalities or accidents is used rather than an absolute number in order to provide a reasonable means of comparison. This is done by dividing the number of deaths or accidents with the total number of registered vehicles in a particular year (i.e. per 10,000 vehicles). The time span of the observations is also dependent on the availability of data and it is imperative to prevent the overlapping of safety measures, as described above. For example, if a safety measure is implemented during the early years in which data are available, the end of the analysis period must not exceed 1988. This is due to the fact that these data are provided only for Peninsular Malaysia and therefore,

the data should be segregated for both Peninsular and East Malaysia until year 1988. Furthermore, it is crucial that there are no major safety measures or interventions that will influence the measured rates during the each period of observation. Table 4.2 shows the type of data used to analyse the impact of each road safety measure implemented in Malaysia.

Table 4.2: Type of data used in univariate analysis and its corresponding period of observations

| Road safety intervention | Type of data used to examine the impact of the intervention | Period of observation |
|---|---|-----------------------|
| Safety Helmet Rules in 1973 | Rate of motorcyclist fatalities | 1969–1988 |
| Seat Belt Rules in 1978 | Rate of motorcar driver fatalities and vulnerable road users fatalities | 1969–1988 |
| Road Transport Act 1987 for Drunken Driving | Rate of accidents due to alcohol | 1969–1996 |
| Speed Limit Rules in 1989 | Rate of motor vehicle accidents | 1969–1994 |
| Motorcycle Daytime Running Headlight Regulation in 1992 | Rate of motorcycle accidents | 1969–1996 |
| Specifications for Protective Helmets (MS 1:1996) | Rate of motorcycle fatalities related to head injuries | 1992–2012 |
| Road Safety Programmes in 1997 | Rate of total road traffic fatalities | 1981–2000 |
| Integrated Road Safety Operations (<i>Ops Sikap</i>) since 2001 | Rate of total road traffic fatalities | 1981–2005 |
| Road Safety Plan 2006–2010 | Rate of total road traffic fatalities | 1981–2012 |

4.3.2 Data series plots

The data series are analysed in the form of graphs in order to determine the appropriate preliminary form of the intervention model for the series, and or determine whether data series are stationary. These plots also offer a number of benefits in the development of the transfer function model.

Firstly, the plots can be used to determine whether the intervention response is a pulse or step function. Secondly, the time plots are very useful to determine the tentative order of the delay time between the response and the intervention applied (b), the shape of the response curve (r) – whether it is exponential or a wave – and the number of transfer functions (ω) before they begin to decay (s) in the transfer function model.

The plot of each data series and the year in which a particular safety measure (intervention) is implemented is shown in Figures 4.20 - 4.29. The details of the data series used in this study are given in Appendix A.

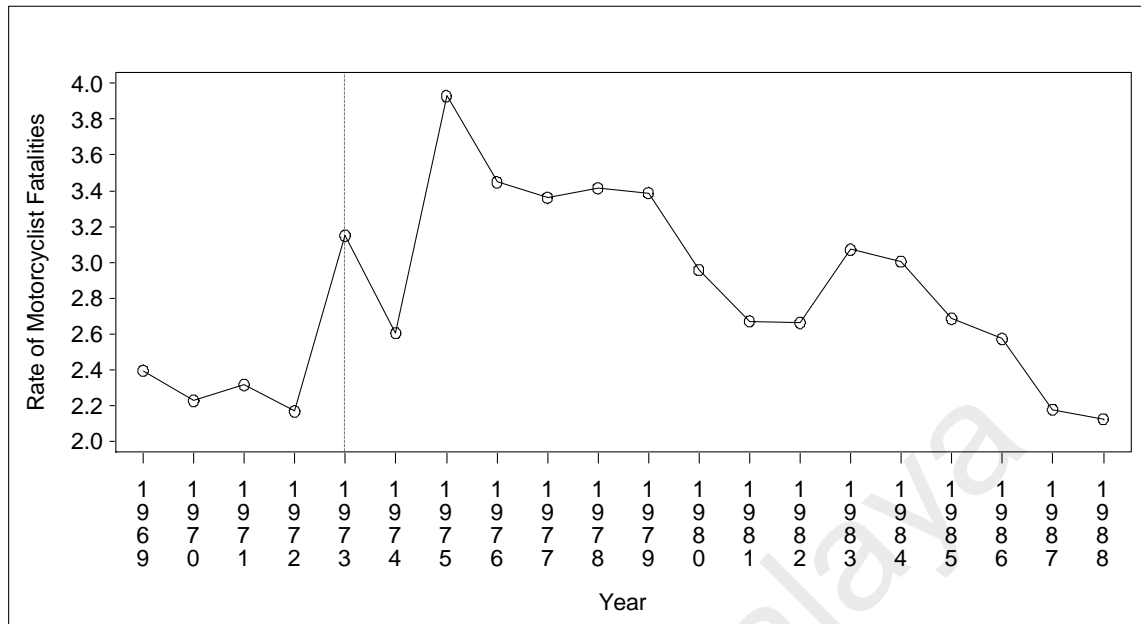


Figure 4.20: Rate of motorcyclist fatalities from 1969 to 1988 and the year in which the safety helmet rules came into force

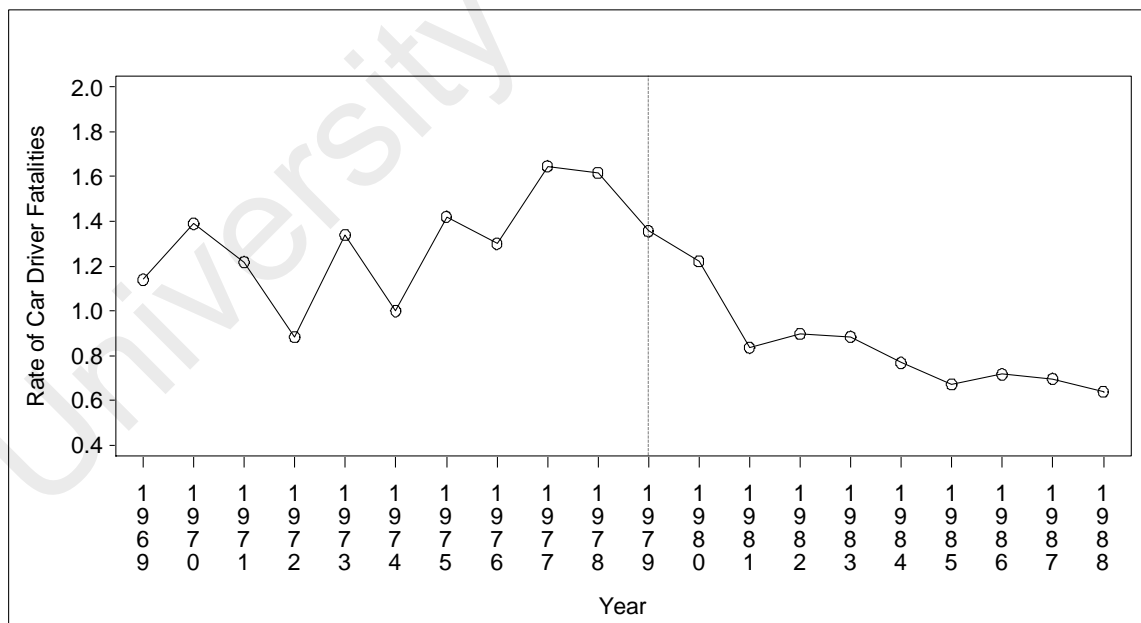


Figure 4.21: Rate of motorcar driver fatalities from 1969 to 1988 and the year in which the seat belt rules came into force

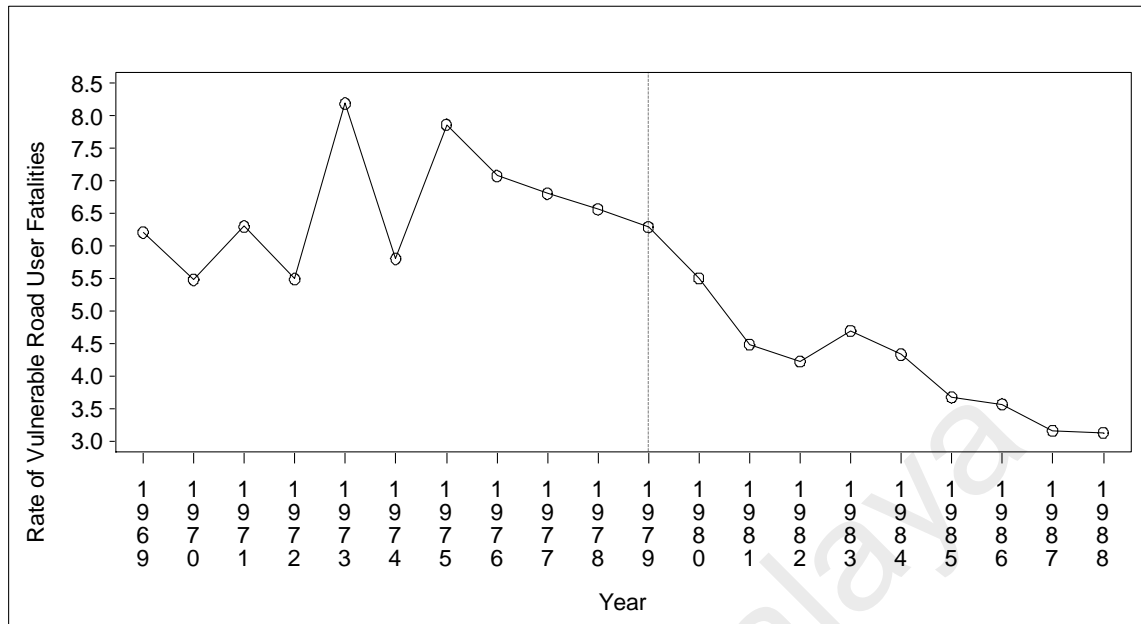


Figure 4.22: Rate of vulnerable road user fatalities from 1969 to 1988 and the year in which the seat belt rules came into force

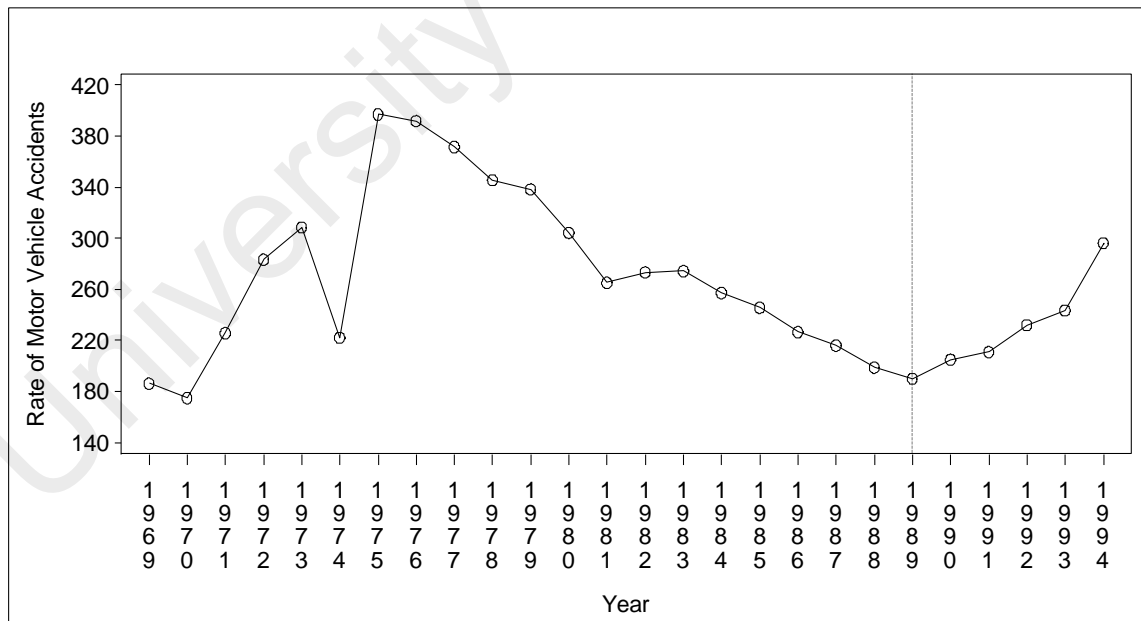


Figure 4.23: Rate of motor vehicle accidents from 1969 to 1994 and the year in which the speed limit rules came into force

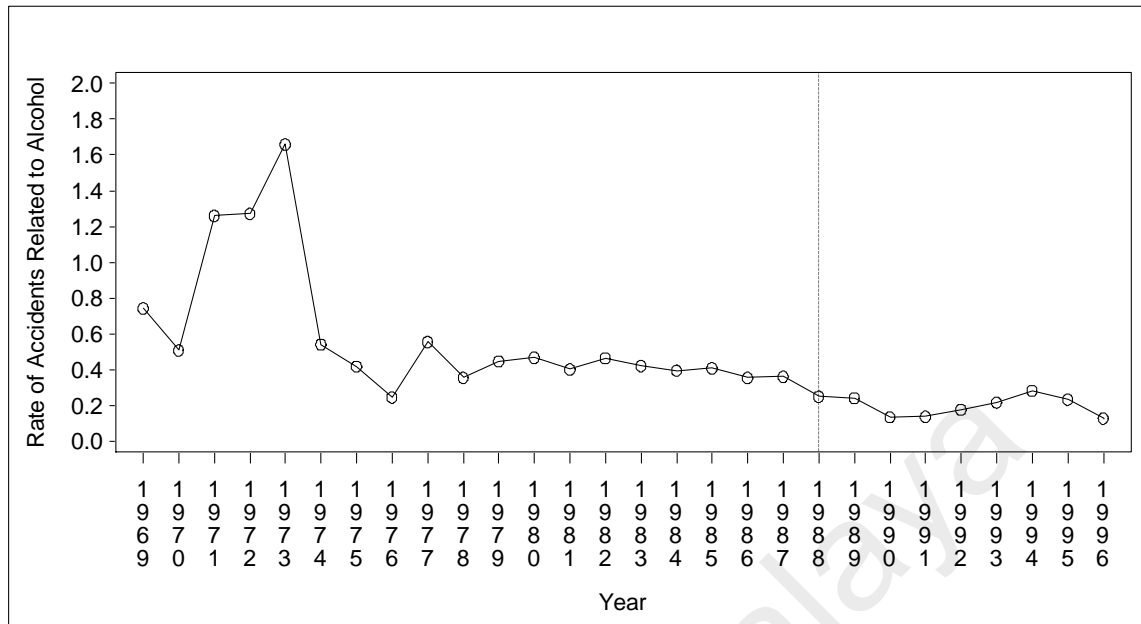


Figure 4.24: Rate of accidents related to alcohol from 1969 to 1996 and the year in which the rules for driving while under the influence of intoxicating liquor or drugs came into force

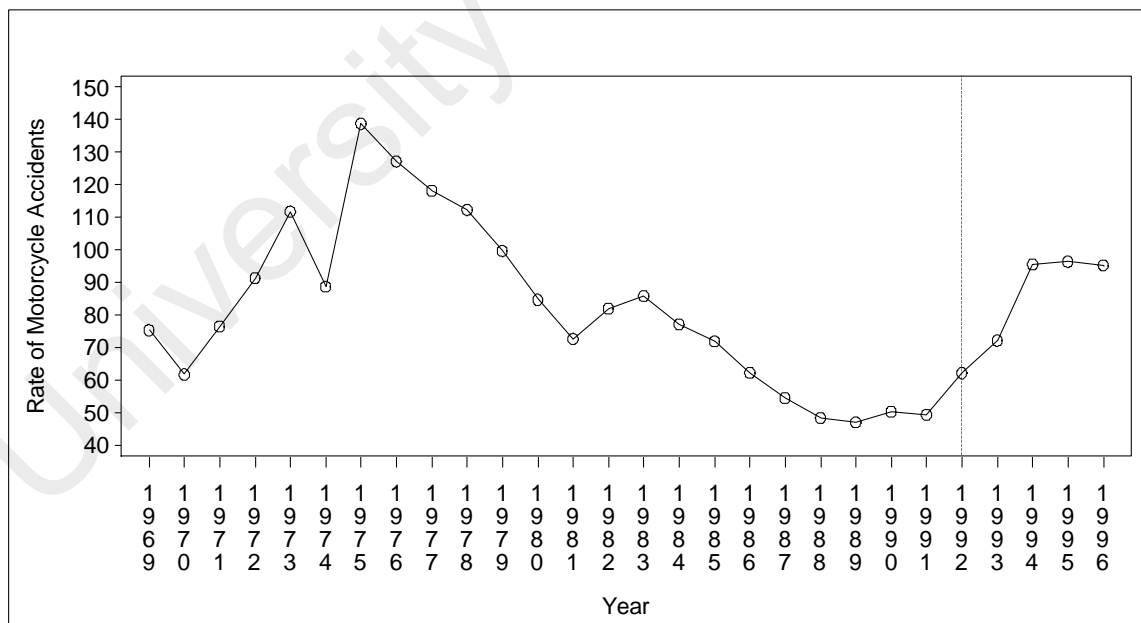


Figure 4.25: Rate of motorcycle accidents from 1969 to 1996 and the year in which the regulation for motorcyclists to switch on headlights during daytime was made compulsory

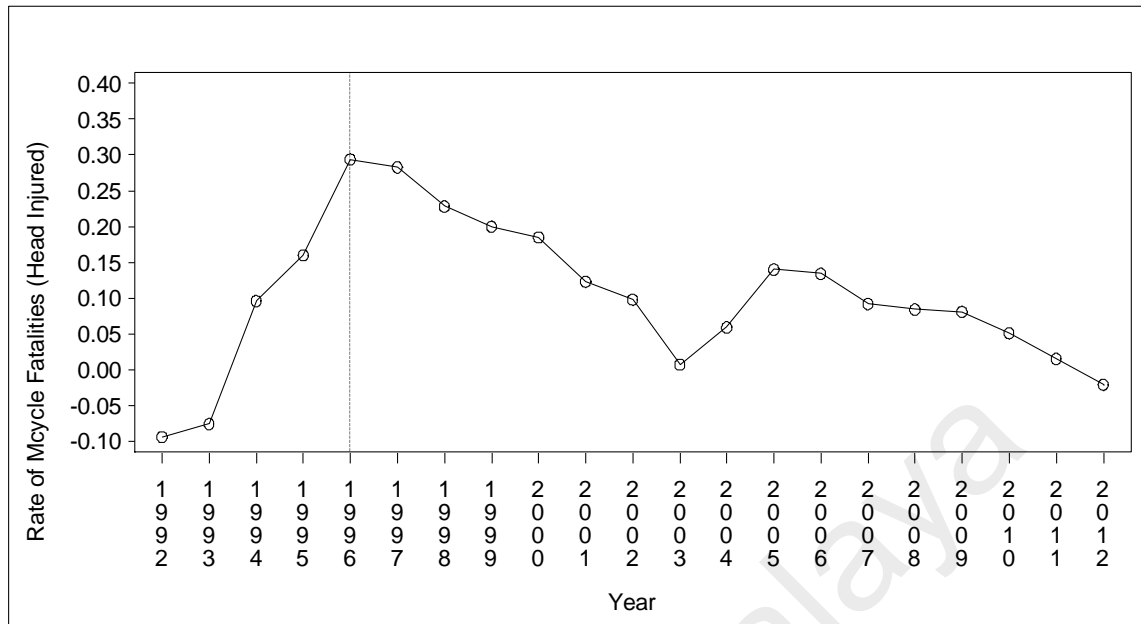


Figure 4.26: Rate of motorcycle fatalities related to head injuries from 1992 to 2012 and the year in which the second revision of the helmet standards was imposed

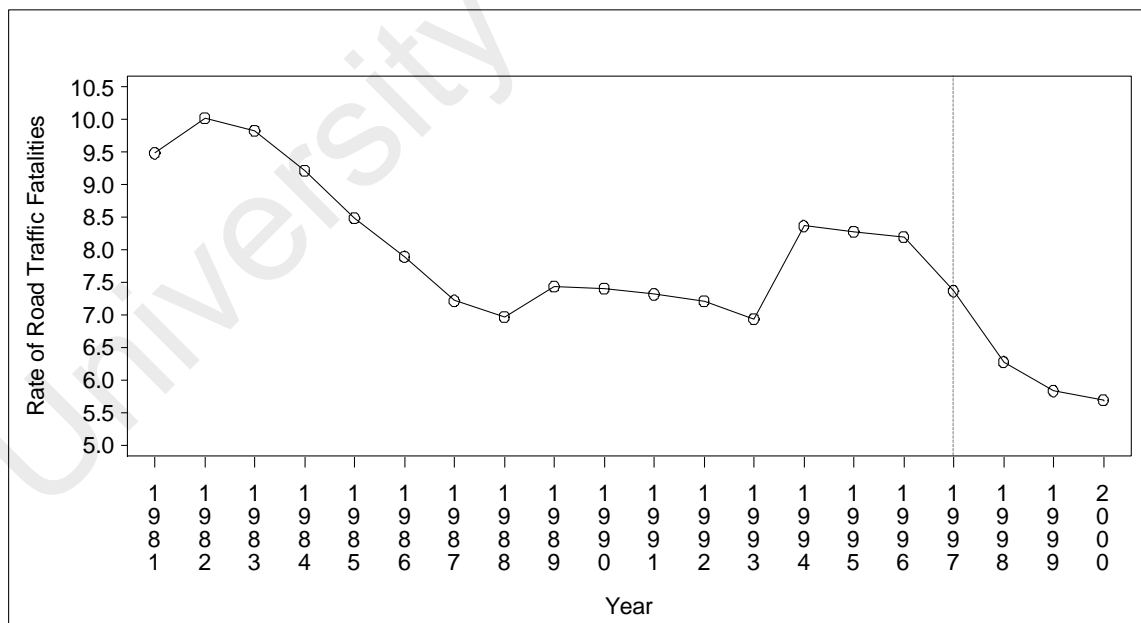


Figure 4.27: Rate of road traffic fatalities from 1981 to 2000 and the year in which comprehensive safety programmes were implemented

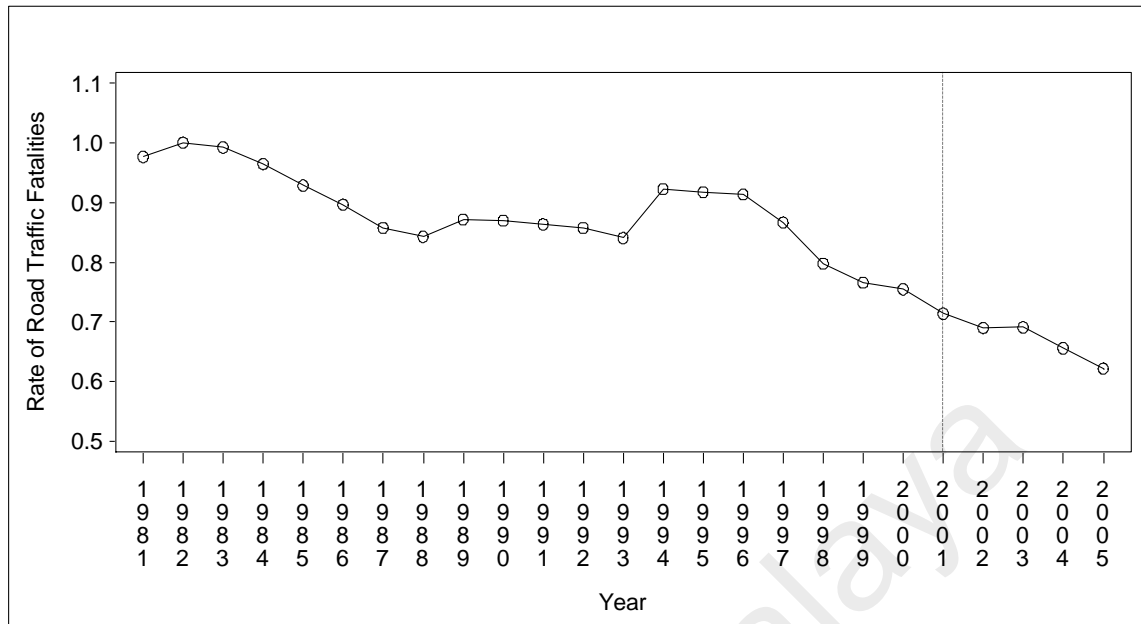


Figure 4.28: Rate of road traffic fatalities from 1981 to 2005 and the year in which the Integrated Road Safety Operations (*Ops Sikap*) was imposed

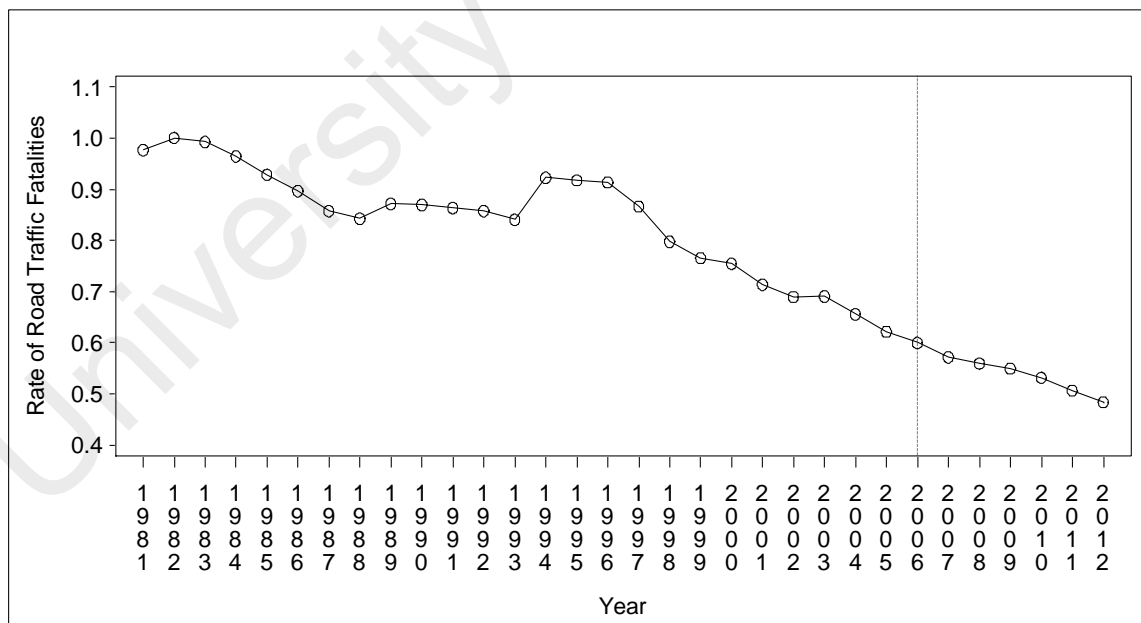


Figure 4.29: Rate of road traffic fatalities from 1981 to 2012 and the year in which the National Road Safety Plan 2006 was launched

4.3.3 ARIMA modelling

The effect of implementing the Integrated Road Safety Operations (*Ops Sikap*) since 2001 on the rate of road traffic fatalities is investigated in order to demonstrate the steps involved in the ARIMA modelling process. The number of fatalities from 1981 to 2000 (a year before the intervention takes place) is used for this purpose, as shown in Table 4.3. The ARIMA model, available in the SAS statistical software, is used to forecast the rate of road traffic fatalities in the absence of *Ops Sikap*. However, it shall be noted that the data must be stationary proceeding with the ARIMA modelling. The SAS programming syntax used to check whether the series is stationary (white noise) is shown in Figure 4.30, whereas the results of the test are shown in Figure 4.31.

It can be seen from Figure 4.31 that the series of fatalities from 1981 to 2000 is not yet stationary. The sample ACF attenuates slowly and the Q -statistic (as shown by the chi-square) of the test for white noise has a p -value less than 0.05. Hence, two methods are used to make the series stationary. Box-Cox transformation is first carried out using Minitab statistical software, followed by the differencing method. The results of Box-Cox transformation (Figure 4.32) show that transformation is required for the series. Logarithmic transformation is used for this purpose. Hence, the log rate of fatalities from 1981 to 2000 is used (Table 4.4) rather than the rate of fatalities. In addition, there is need to determine whether the data series is stationary after log transformation. The SAS outputs are shown Figure 4.33, which indicate that the log rate of fatalities is not stationary (white noise).

The next step involves determining the order of the differencing required using the Augmented Dickey-Fuller (ADF) test. The SAS statements used for the ADF test are shown in Figure 4.34, and the results of the ADF test are shown in Figure 4.35 and 4.36. Based on Figure 4.28, the rate of road traffic fatalities can be classified as a random walk

series with/without a trend. Thus, the values of the ADF test with a single mean and trend used as the basis to determine whether the data series is stationary. Figure 4.36 shows that the data series is stationary upon application of first-order differencing.

Table 4.3: Rate of road traffic fatalities from 1981 to 2000

| Year | Rate of road traffic fatalities | Year | Rate of road traffic fatalities | Year | Rate of road traffic fatalities |
|------|---------------------------------|------|---------------------------------|------|---------------------------------|
| 1981 | 9.487 | 1988 | 6.973 | 1995 | 8.281 |
| 1982 | 10.026 | 1989 | 7.439 | 1996 | 8.201 |
| 1983 | 9.834 | 1990 | 7.410 | 1997 | 7.370 |
| 1984 | 9.219 | 1991 | 7.319 | 1998 | 6.279 |
| 1985 | 8.490 | 1992 | 7.215 | 1999 | 5.835 |
| 1986 | 7.896 | 1993 | 6.941 | 2000 | 5.694 |
| 1987 | 7.218 | 1994 | 8.366 | | |

```

data safety;
    input year fatalities;
    datalines;
1981      9.487
1982     10.026
1983      9.834
1984      9.219
1985      8.490
1986      7.896
1987      7.218
1988      6.973
1989      7.439
1990      7.410
.....
2000      5.694
;

proc arima data=safety;
    identify var=fatalities nlag=10;
run;

```

Figure 4.30: SAS syntax used to check stationarity of the rate of fatalities series from 1981 to 2000

Table 4.4: Log rate of road traffic fatalities from 1981 to 2000

| Year | Log rate of fatalities | Year | Log rate of fatalities | Year | Log rate of fatalities |
|------|------------------------|------|------------------------|------|------------------------|
| 1981 | 0.977 | 1988 | 0.843 | 1995 | 0.918 |
| 1982 | 1.001 | 1989 | 0.872 | 1996 | 0.914 |
| 1983 | 0.993 | 1990 | 0.870 | 1997 | 0.867 |
| 1984 | 0.965 | 1991 | 0.864 | 1998 | 0.798 |
| 1985 | 0.929 | 1992 | 0.858 | 1999 | 0.766 |
| 1986 | 0.897 | 1993 | 0.841 | 2000 | 0.755 |
| 1987 | 0.858 | 1994 | 0.923 | | |

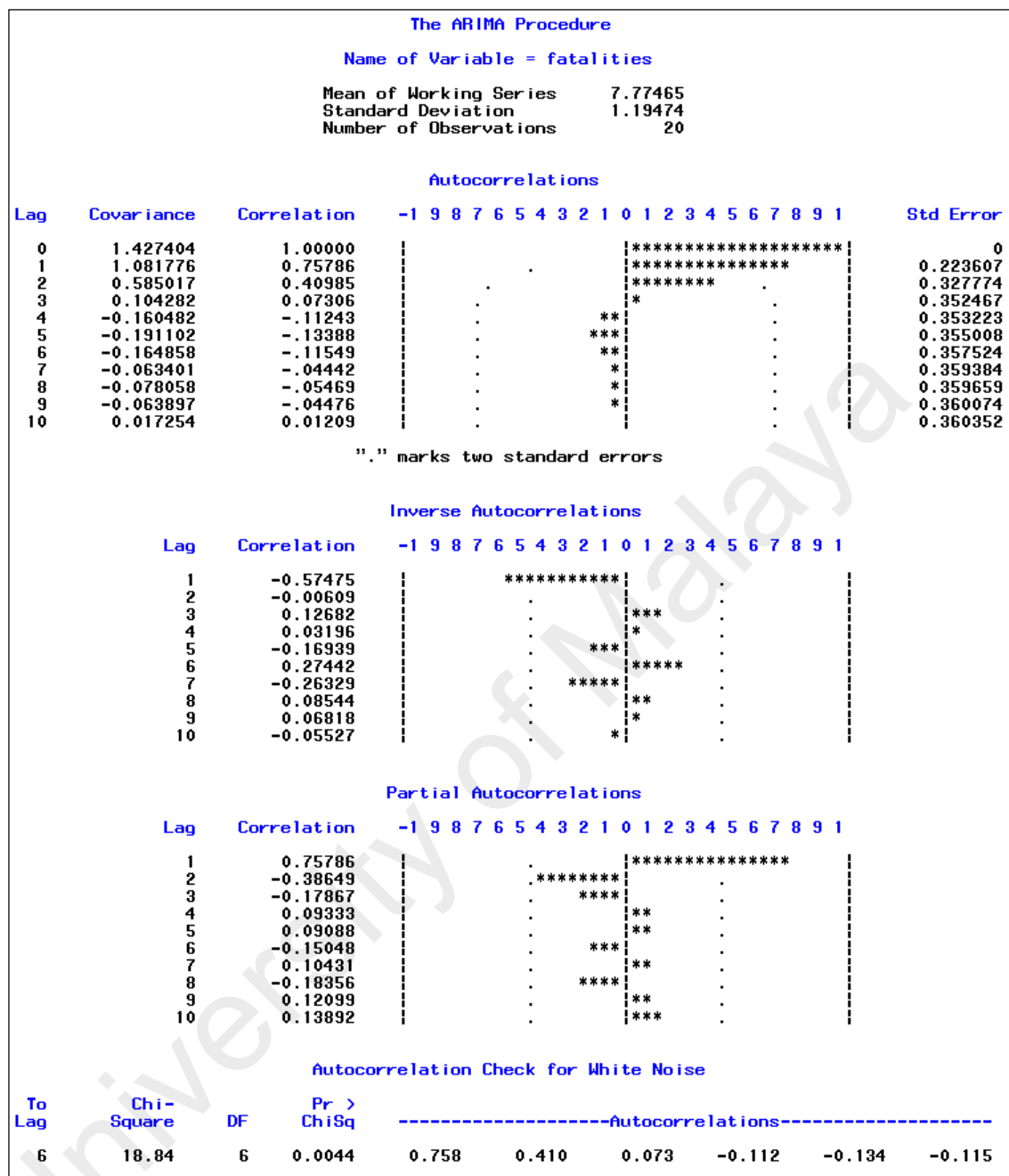


Figure 4.31: Sample ACF, IACF, PACF and autocorrelation check for white noise of the log rate of fatalities series from 1981 to 2000

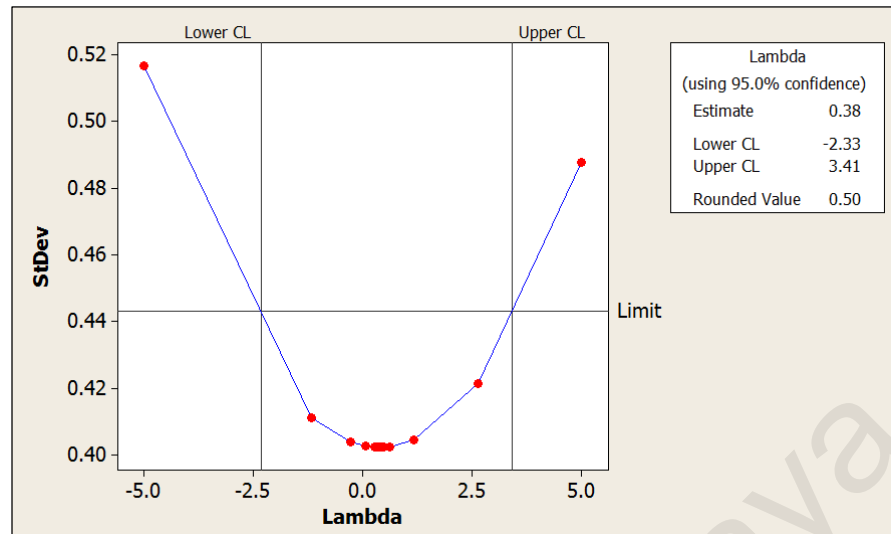


Figure 4.32: Lambda (λ) value of the log rate of fatalities from 1981 to 2000

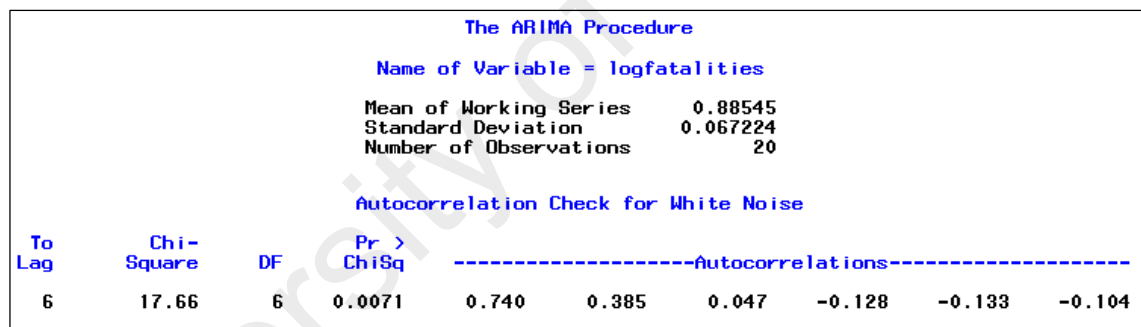


Figure 4.33: Results of the autocorrelation check for white noise for the log rate of fatalities from 1981 to 2000

Since stationarity is attained for the series using log transformation and first-order differencing, the Box-Jenkins method is used to develop the ARIMA model that is most suitable to forecast the log rate of fatalities. The maximum likelihood method is chosen to estimate the parameters of the model.

```

data safety;
    input year logfatalities;
    datalines;
1981    0.977
1982    1.001
1983    0.993
1984    0.965
1985    0.929
1986    0.897
1987    0.858
1988    0.843
1989    0.872
1990    0.870
.....
2000    0.755
;

proc arima data=safety;
    identify var=logfatalities /* Test of No Differenced for the series */
    nlag=7 stationarity=(adf=(0,1,2,3,4,5,6,7));
    /* Augmented Dickey-Fuller Tests */
    title1 'Augmented Dickey-Fuller Tests'; /* at lags 0 through 7 */
run;

```

Figure 4.34: SAS syntax used to conduct the Augmented Dickey-Fuller (ADF) test for the log rate of fatalities from 1981 to 2000

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|------|----------|----------|-------|----------|------|--------|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | -0.2538 | 0.6096 | -1.61 | 0.0987 | | |
| | 1 | -0.3105 | 0.5970 | -1.43 | 0.1351 | | |
| | 2 | -0.2866 | 0.6006 | -1.25 | 0.1847 | | |
| | 3 | -0.2049 | 0.6183 | -1.16 | 0.2132 | | |
| | 4 | -0.1536 | 0.6289 | -1.27 | 0.1765 | | |
| | 5 | -0.1498 | 0.6293 | -0.91 | 0.3027 | | |
| | 6 | -0.1023 | 0.6375 | -0.74 | 0.3743 | | |
| | 7 | -0.0998 | 0.6373 | -0.88 | 0.3150 | | |
| Single Mean | 0 | -1.2967 | 0.8419 | -0.55 | 0.8595 | 1.34 | 0.7417 |
| | 1 | -6.8393 | 0.2280 | -1.68 | 0.4238 | 2.39 | 0.4998 |
| | 2 | -13.6273 | 0.0192 | -1.88 | 0.3329 | 2.54 | 0.4637 |
| | 3 | -8.6994 | 0.1205 | -1.39 | 0.5578 | 1.61 | 0.6791 |
| | 4 | -5.2896 | 0.3468 | -1.21 | 0.6390 | 1.51 | 0.7021 |
| | 5 | -23.1432 | <.0001 | -1.44 | 0.5302 | 1.46 | 0.7127 |
| | 6 | -12.5352 | 0.0206 | -0.87 | 0.7648 | 0.62 | 0.9087 |
| | 7 | -2.1177 | 0.7322 | -0.08 | 0.9307 | 0.29 | 0.9899 |
| Trend | 0 | -4.5715 | 0.8194 | -1.37 | 0.8362 | 1.06 | 0.9524 |
| | 1 | -15.4904 | 0.0621 | -2.42 | 0.3588 | 2.93 | 0.6211 |
| | 2 | -128.189 | 0.0001 | -2.99 | 0.1638 | 4.46 | 0.3539 |
| | 3 | 101.2011 | 0.9999 | -2.41 | 0.3619 | 2.90 | 0.6259 |
| | 4 | -305.761 | 0.0001 | -1.71 | 0.6985 | 1.48 | 0.8745 |
| | 5 | 10.4283 | 0.9999 | -2.33 | 0.3931 | 2.71 | 0.6582 |
| | 6 | 5.0218 | 0.9999 | -2.13 | 0.4869 | 2.49 | 0.6980 |
| | 7 | 1.4327 | 0.9955 | -0.30 | 0.9779 | 6.07 | 0.0942 |

Figure 4.35: Results of the ADF test for log rate of fatalities without differencing

| The ARIMA Procedure | | | | | | | | | |
|---|------------|----------|------------|----------------------------|----------|--------|--------|--------|--------|
| Name of Variable = logfatalities | | | | | | | | | |
| Period(s) of Differencing | | | | 1 | | | | | |
| Mean of Working Series | | | | -0.01168 | | | | | |
| Standard Deviation | | | | 0.031608 | | | | | |
| Number of Observations | | | | 19 | | | | | |
| Observation(s) eliminated by differencing | | | | 1 | | | | | |
| Autocorrelation Check for White Noise | | | | | | | | | |
| To Lag | Chi-Square | DF | Pr > ChiSq | -----Autocorrelations----- | | | | | |
| 6 | 6.41 | 6 | 0.3789 | 0.264 | 0.038 | -0.247 | -0.351 | -0.001 | -0.029 |
| Dickey-Fuller Unit Root Tests | | | | | | | | | |
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F | | |
| Zero Mean | 0 | -11.3814 | 0.0098 | -2.84 | 0.0071 | | | | |
| Single Mean | 0 | -13.2535 | 0.0242 | -3.18 | 0.0386 | 5.08 | 0.0548 | | |
| Trend | 0 | -13.2108 | 0.1324 | -3.04 | 0.1492 | 4.73 | 0.3053 | | |

Figure 4.36: Results of the ADF test and autocorrelation check for white noise for the log rate of fatalities series with first-order differencing

The next step involved is to determine the tentative order of the ARIMA model. In order to generate various potential ARIMA models that are representative of the series, the $AR(p)$ and $MA(q)$ are included in the model identification, in which the order ranges from 0 to 12. The SAS statements used to generate the model identification is shown in Figure 4.37, and a total of 169 different model equations are produced by the software. The p -value of model's parameters is the first criterion used to judge whether a model is sufficient before proceeding to the next step. In this case, the critical p -value is 0.1. It can be seen from Table 4.5 that there are 10 models (out of the 169 models produced) in which the p -values of all parameters are less than 0.1. These models are then examined to determine the appropriate models. The information criteria (i.e. AIC and SBC) are used as second criterion to screen the models, and the results are shown in Table 4.5. In this study, the model that fits best according to the Schwartz criterion (i.e. SBC or BIC) is adopted. The model is selected based on the lowest SBC in order to obtain a simple model.

Referring to Table 4.5, it can be seen that the ARIMA(12,1,4) model has the lowest SBC value, i.e. -77.67. Thus, this model is tentatively considered as the appropriate

model. However, model adequacy checks need to be carried out in order to confirm that the model is stable and adequate. This is done by checking the Q -statistic as well as the normality of the model with the lowest SBC. The Kolmogorov-Smirnov p -value is adopted for the normality test and is used to determine whether the forecast is normally distributed or not. The SAS statements used to complete both tests are shown in Figure 4.38.

```

data safety;
    input year logfatalities;
    datalines;
1981    0.977
1982    1.001
1983    0.993
1984    0.965
1985    0.929
1986    0.897
1987    0.858
1988    0.843
1989    0.872
1990    0.870
.....
2000    0.755
;

proc arima data=safety;
    i var=logfatalities(1) noprint;
    /*--- (1) data first differencing nonseasonal d=1 ---*/
    e p=(0) q=(0) method=ml;
run;

    e p=(0) q=(1) method=ml;
    e p=(0) q=(2) method=ml;
    .....
    e p=(6) q=(10) method=ml;
    e p=(6) q=(11) method=ml;
    e p=(6) q=(12) method=ml;
    .....
    e p=(12) q=(10) method=ml;
    e p=(12) q=(11) method=ml;
    e p=(12) q=(12) method=ml;
run;

```

Figure 4.37: SAS programming syntax used to determine the potential ARIMA(p,d,q) models which will forecast the log rate of fatalities from 1981 to 2000

Table 4.5: ARIMA models in which the parameters' p -values are less than 0.1 and their information criteria

| No. | ARIMA (p,d,q) | AIC | SBC | No. | ARIMA (p,d,q) | AIC | SBC |
|-----|-------------------|--------|--------|-----|-------------------|--------|--------|
| 1 | ARIMA(12,1,4) | -80.51 | -77.67 | 6 | ARIMA(0,1,3) | -75.44 | -73.55 |
| 2 | ARIMA(4,1,3) | -79.69 | -76.85 | 7 | ARIMA(9,1,7) | -76.17 | -73.33 |
| 3 | ARIMA(12,1,8) | -79.56 | -76.73 | 8 | ARIMA(7,1,9) | -75.30 | -72.46 |
| 4 | ARIMA(0,1,4) | -77.72 | -75.84 | 9 | ARIMA(11,1,9) | -73.82 | -70.98 |
| 5 | ARIMA(0,1,8) | -77.00 | -75.12 | 10 | ARIMA(7,1,7) | -72.29 | -69.46 |

The Q -statistic and the Kolmogorov-Smirnov p -value of the log rate of fatalities from 1981 to 2000 are shown in Figure 4.39. The results of the adequacy check of the ARIMA(12,1,4) model verifies that the model is stable and appropriate. It can be seen that the Q -statistic (chi-square) at all lags and the p -values of the Kolmogorov-Smirnov normality test for the model are greater than 0.05. Hence, it can be concluded that the ARIMA(12,1,4) model is appropriate to forecast the series. Based on the SAS outputs shown in Figure 4.39, the mathematical representation of the ARIMA(12,1,4) including each parameter's p -value (indicated within the parentheses) is given by:

$$(0.0060)$$

$$(1 - B)Y_t = -0.00834 + \frac{(1 - 0.51466B^4)}{(1 - 0.67104B^{12})} a_t \quad (4.1)$$

(0.0274) (0.0467)

```

data safety;
    input year logfatalities;
    datalines;
1981      0.977
1982      1.001
1983      0.993
1984      0.965
1985      0.929
1986      0.897
1987      0.858
1988      0.843
1989      0.872
1990      0.870
.....
2000      0.755
;

proc arima data=safety;
    i var=logfatalities(1) noprint;
    /*--- (1) data first differencing nonseasonal d=1 ---*/
    e p=(12) q=(4) method=ml;
run;
forecast lead=5 out=fore
    id=year interval=1 printall;
run;

proc arima data=fore;
    identify var=residual;
run;

proc univariate data=fore normal plot; /*--- normality test ---*/
    var residual;
run;

```

Figure 4.38: SAS syntax used to check the adequacy of the ARIMA(12,1,4) model adequacy forecast the fatalities

| The ARIMA Procedure | | | | | | | | | |
|------------------------------------|---------------|----------------|------------|----------------------------|--------|--------|--------|--------|-------|
| Maximum Likelihood Estimation | | | | | | | | | |
| Parameter | Estimate | Standard Error | t Value | Approx Pr > t | Lag | | | | |
| MU | -0.0083435 | 0.0037821 | -2.21 | 0.0274 | 0 | | | | |
| MA1,1 | 0.51466 | 0.18727 | 2.75 | 0.0060 | 4 | | | | |
| AR1,1 | 0.67104 | 0.33735 | 1.99 | 0.0467 | 12 | | | | |
| Constant Estimate | | | -0.00274 | | | | | | |
| Variance Estimate | | | 0.000491 | | | | | | |
| Std Error Estimate | | | 0.02216 | | | | | | |
| AIC | | | -80.5058 | | | | | | |
| SBC | | | -77.6725 | | | | | | |
| Number of Residuals | | | 19 | | | | | | |
| Autocorrelation Check of Residuals | | | | | | | | | |
| To Lag | Chi-Square | DF | Pr > ChiSq | -----Autocorrelations----- | | | | | |
| 6 | 3.31 | 4 | 0.5073 | 0.062 | -0.056 | -0.347 | -0.024 | 0.058 | 0.032 |
| 12 | 9.51 | 10 | 0.4841 | -0.056 | -0.098 | -0.249 | -0.083 | 0.050 | 0.224 |
| 18 | 10.65 | 16 | 0.8304 | 0.030 | 0.053 | -0.064 | -0.041 | -0.012 | 0.022 |
| Model for variable logfatalities | | | | | | | | | |
| Estimated Mean | | | -0.00834 | | | | | | |
| Period(s) of Differencing | | | 1 | | | | | | |
| Autoregressive Factors | | | | | | | | | |
| Factor 1: 1 - 0.67104 B**(12) | | | | | | | | | |
| Moving Average Factors | | | | | | | | | |
| Factor 1: 1 - 0.51466 B**(4) | | | | | | | | | |
| Tests for Normality | | | | | | | | | |
| Test | --Statistic-- | | | -----p Value----- | | | | | |
| Kolmogorov-Smirnov | D | 0.170566 | Pr > D | 0.1500 | | | | | |

Figure 4.39: Parameters, residuals and adequacy check results of the ARIMA(12,1,4) model

The following step involved is to attain the forecasted values of the log rate of fatalities using the ARIMA(12,1,4) model and the results are shown in Figure 4.40. The forecast period is from 1981 and extended to 2005. This is to fulfil the two objectives: (1) to compare the forecasted log rate of fatalities obtained from the model with the actual ones (Figure 4.41) and (2) to attain the log rate of fatalities after year 2000. The log rate of fatalities beyond year 2000 obtained from the ARIMA model shows the rate of fatalities if there no intervention (*Ops Sikap*) was implemented since 2001 onwards. It can be noticed that the rates of fatalities tend to increase beyond year 2000. These values will be

compared with the forecasted values obtained from the intervention model, as described in the following section.

| The ARIMA Procedure | | | | | | |
|--------------------------------------|----------|-----------|-----------------------|--------|--------|----------|
| Forecasts for variable logfatalities | | | | | | |
| Obs | Forecast | Std Error | 95% Confidence Limits | | Actual | Residual |
| 2 | 0.9687 | 0.0336 | 0.9028 | 1.0345 | 1.0010 | 0.0323 |
| 3 | 0.9927 | 0.0336 | 0.9268 | 1.0585 | 0.9930 | 0.0003 |
| 4 | 0.9847 | 0.0336 | 0.9188 | 1.0505 | 0.9650 | -0.0197 |
| 5 | 0.9567 | 0.0336 | 0.8908 | 1.0225 | 0.9290 | -0.0277 |
| 6 | 0.9075 | 0.0307 | 0.8473 | 0.9677 | 0.8970 | -0.0105 |
| 7 | 0.8885 | 0.0307 | 0.8283 | 0.9487 | 0.8580 | -0.0305 |
| 8 | 0.8577 | 0.0307 | 0.7975 | 0.9178 | 0.8430 | -0.0147 |
| 9 | 0.8459 | 0.0307 | 0.7857 | 0.9061 | 0.8720 | 0.0261 |
| 10 | 0.8613 | 0.0261 | 0.8101 | 0.9125 | 0.8700 | 0.0087 |
| 11 | 0.8805 | 0.0261 | 0.8293 | 0.9317 | 0.8640 | -0.0165 |
| 12 | 0.8701 | 0.0261 | 0.8189 | 0.9213 | 0.8580 | -0.0121 |
| 13 | 0.8410 | 0.0261 | 0.7898 | 0.8922 | 0.8410 | -0.0000 |
| 14 | 0.8512 | 0.0230 | 0.8061 | 0.8962 | 0.9230 | 0.0718 |
| 15 | 0.9210 | 0.0230 | 0.8760 | 0.9660 | 0.9180 | -0.0030 |
| 16 | 0.9010 | 0.0230 | 0.8559 | 0.9460 | 0.9140 | 0.0130 |
| 17 | 0.8871 | 0.0230 | 0.8421 | 0.9321 | 0.8670 | -0.0201 |
| 18 | 0.8084 | 0.0224 | 0.7645 | 0.8522 | 0.7980 | -0.0104 |
| 19 | 0.7705 | 0.0224 | 0.7267 | 0.8144 | 0.7660 | -0.0045 |
| 20 | 0.7469 | 0.0224 | 0.7031 | 0.7908 | 0.7550 | 0.0081 |
| 21 | 0.7813 | 0.0222 | 0.7379 | 0.8248 | . | . |
| 22 | 0.7825 | 0.0313 | 0.7211 | 0.8439 | . | . |
| 23 | 0.7780 | 0.0384 | 0.7028 | 0.8532 | . | . |
| 24 | 0.7672 | 0.0443 | 0.6803 | 0.8540 | . | . |
| 25 | 0.7530 | 0.0456 | 0.6636 | 0.8424 | . | . |

Figure 4.40: Forecasted log rate of fatalities produced by the ARIMA(12,1,4) model

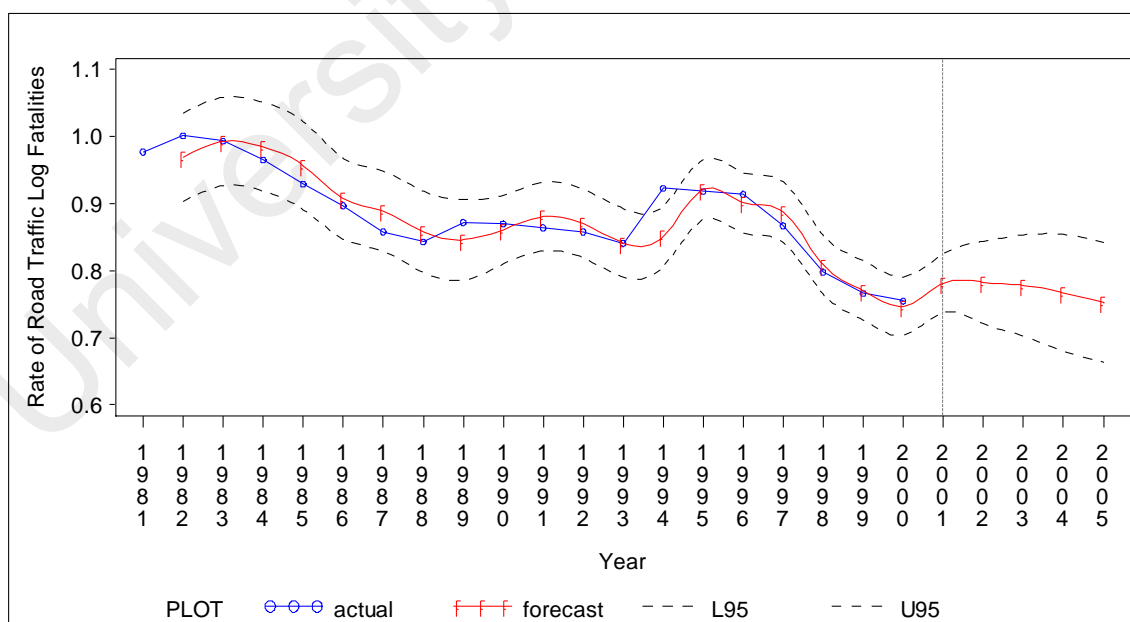


Figure 4.41: Time plot of actual and forecasted log rate of fatalities produced by the ARIMA(12,1,4) model

4.3.4 Intervention analysis modelling

The time plots of the rate of injuries or accidents shown in Figures 4.20 through 4.29 are examined in detail, and in summary, the interventions (safety measures) implemented have led to changes in the trend of the rate of casualties or accidents. Further analysis on the trend of the rate of fatalities leads to the conclusion that the transfer function-noise model is suitable for the intervention analysis. To provide a continuity of the ARIMA modelling above, the effect of implementing the Integrated Road Safety Operations (*Ops Sikap*) since 2001 on the rate of road traffic fatalities is used as an example to demonstrate the steps taken in the modelling process. The first step involved is to determine whether the series is autocorrelated and eliminate any autocorrelations to provide a white noise series. This should be done since the ARIMA model uses the data series from 1981 to 2000. However, in series used in intervention modelling is from 1981 to 2005. The SAS programming syntax used to complete this task is shown in Figure 4.42, while the results are shown in Figure 4.43.

```
data safety;
    input year fatalities;
    datalines;
1981      9.487
1982     10.026
1983      9.834
1984      9.219
1985      8.490
1986      7.896
1987      7.218
1988      6.973
1989      7.439
1990      7.410
.....
2005      4.185
;

proc arima data=safety;
    identify var=fatalities nlag=10;
run;
```

Figure 4.42: SAS syntax used to check stationarity in the rate of fatalities from 1981 to 2005

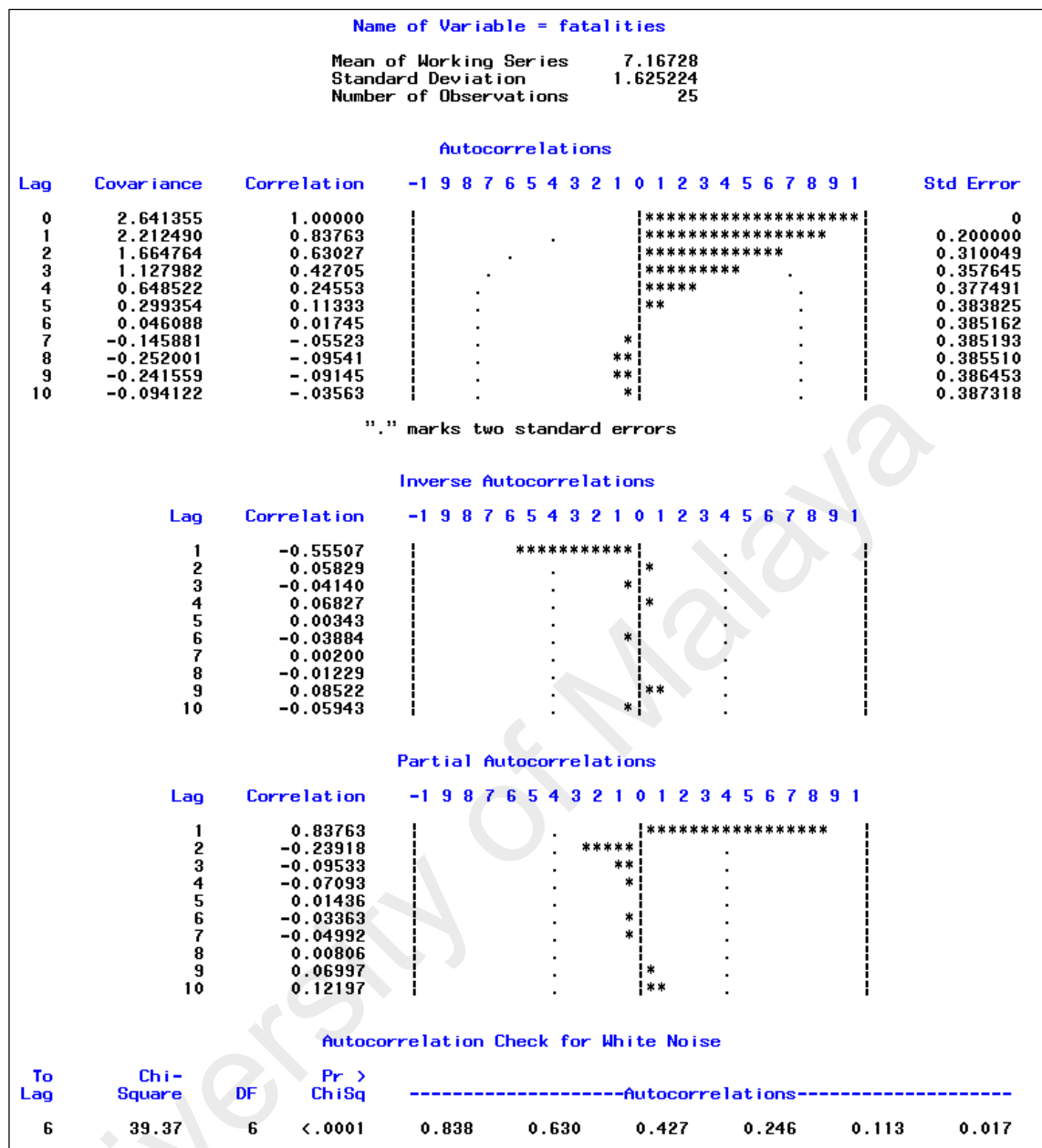


Figure 4.43: Sample ACF, IACF, PACF and autocorrelation check for white noise in rate of fatalities from 1981 to 2005

It can be seen from the sample ACF and PACF pattern in Figure 4.43 that rate of fatalities is not yet stationary. This is verified based on the p -value of the chi-square statistic, which is less than 0.05. Hence, there is need to pre-whiten the series. The Box-Cox transformation method is initially carried out to check and transform the data in order to achieve stationarity. The lambda (λ) value becomes the basis during data

transformation. Logarithm transformation is, however, prioritised in transforming the data. Other equations from the Box-Cox method will be used if the data are not stationary after the application of logarithm transformation. Stationarity in variance is achieved by applying the logarithm and/or other Box-Cox transformation, whereas differencing method will be used to achieve stationarity in mean. The unit root test is then executed using the Augmented Dickey-Fuller (ADF) test to determine whether the data are stationary. Differencing method will be applied to ensure that stationarity is fulfilled.

Minitab statistical software is used to complete this transforming task. The results of Box-Cox transformation analysis are shown in Figure 4.44. The value of $\lambda = -0.23$ is rounded to 0.00. This indicates that logarithmic equation is suitable to transform the data to attain stationarity in variance. For this reason, the logarithmic equation is used to transform the rate of road traffic fatalities and the results are presented in Table 4.6.

The next step involves checking whether the series is stationary after log transformation of the series. However, the SAS outputs shown in Figure 4.45 indicate that the log rate of fatalities is not yet stationary after log transformation. Autocorrelation check on the log rate of fatalities for white noise reveals that the p -value is less than 0.05. Hence, the differencing method is applied on the series, beginning from first-order differencing.

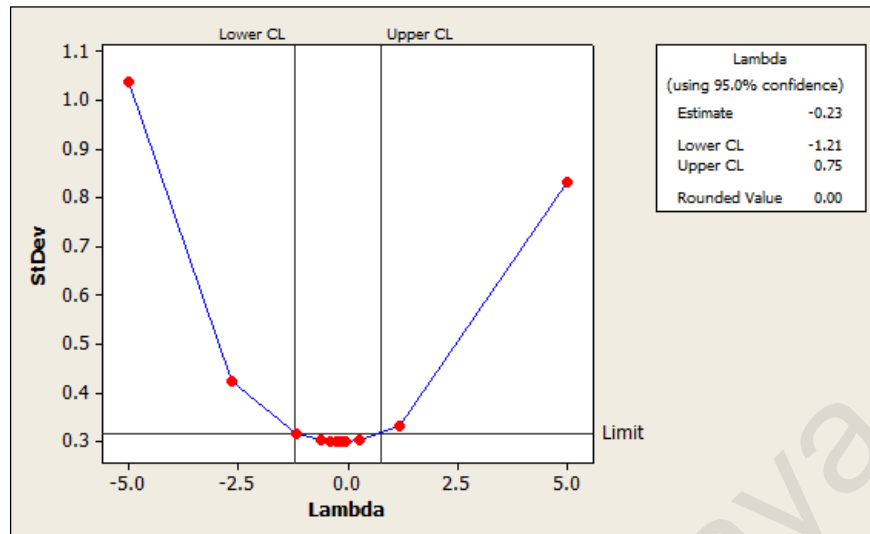


Figure 4.44: Lambda (λ) value of the rate of fatalities data from 1981 to 2005

Table 4.6: Log rate of road traffic fatalities from 1981 to 2005

| Year | Log rate of road traffic fatalities | Year | Log rate of road traffic fatalities | Year | Log rate of road traffic fatalities |
|------|-------------------------------------|------|-------------------------------------|------|-------------------------------------|
| 1981 | 0.977 | 1990 | 0.870 | 1999 | 0.766 |
| 1982 | 1.001 | 1991 | 0.864 | 2000 | 0.755 |
| 1983 | 0.993 | 1992 | 0.858 | 2001 | 0.714 |
| 1984 | 0.965 | 1993 | 0.841 | 2002 | 0.690 |
| 1985 | 0.929 | 1994 | 0.923 | 2003 | 0.691 |
| 1986 | 0.897 | 1995 | 0.918 | 2004 | 0.656 |
| 1987 | 0.858 | 1996 | 0.914 | 2005 | 0.622 |
| 1988 | 0.843 | 1997 | 0.867 | | |
| 1989 | 0.872 | 1998 | 0.798 | | |

| | | | | | | | | | |
|---------------------------------------|----------------|----|---------------|----------------------------|-------|-------|-------|-------|-------|
| Name of Variable = logfatalities | | | | | | | | | |
| Mean of Working Series | | | | 0.84328 | | | | | |
| Standard Deviation | | | | 0.104572 | | | | | |
| Number of Observations | | | | 25 | | | | | |
| Autocorrelation Check for White Noise | | | | | | | | | |
| To Lag | Chi- Square | DF | Pr > ChiSq | -----Autocorrelations----- | | | | | |
| 6 | 40.86 | 6 | <.0001 | 0.830 | 0.637 | 0.456 | 0.278 | 0.140 | 0.038 |

Figure 4.45: Autocorrelation check on the log rate of fatalities for white noise

The ADF test is then conducted to determine the order of the differencing required, and the SAS syntax used for this test is shown in Figure 4.46. The results are shown in Figure 4.47 and Figure 4.48. The results shown in Figure 4.47, however, suggest that differencing is required in order to transform the data series (log rate of fatalities) and make it stationary. Thus, first-order and then second-order differencing (if necessary) is applied to the series. However, the ADF test results indicate that the series is stationary after application of first-order differencing. The ADF test results following first order differencing are presented in Figure 4.48.

```

data safety;
  input year logfatalities;
  datalines;
1981 0.977
1982 1.001
1983 0.993
1984 0.965
1985 0.929
1986 0.897
1987 0.858
1988 0.843
1989 0.872
1990 0.870
.....
2005 0.622
;

proc arima data=safety;
  identify var=logfatalities /* Test of No Differenced for the series */
  nlag=10 stationarity=(adf=(0,1,2,3,4,5,6,7,8,9,10));
  /* Augmented Dickey-Fuller Tests */
  title1 'Augmented Dickey-Fuller Tests'; /* at lags 0 through 10 */
run;

```

Figure 4.46: SAS syntax for the Augmented Dickey-Fuller (ADF) test of log rate of fatalities

| Type | Lags | Rho | <u>Pr</u> < Rho | Tau | <u>Pr</u> < Tau | F | <u>Pr</u> > F |
|-------------|------|----------|-----------------|-------|-----------------|------|---------------|
| Zero Mean | 0 | -0.5839 | 0.5455 | -3.09 | 0.0030 | | |
| | 1 | -0.6706 | 0.5273 | -2.42 | 0.0171 | | |
| | 2 | -0.6522 | 0.5304 | -2.13 | 0.0340 | | |
| | 3 | -0.5962 | 0.5411 | -1.89 | 0.0570 | | |
| | 4 | -0.5395 | 0.5518 | -1.83 | 0.0646 | | |
| | 5 | -0.5305 | 0.5537 | -1.39 | 0.1486 | | |
| | 6 | -0.4633 | 0.5669 | -1.07 | 0.2473 | | |
| | 7 | -0.4400 | 0.5717 | -1.13 | 0.2273 | | |
| | 8 | -0.5166 | 0.5560 | -1.47 | 0.1285 | | |
| | 9 | -0.5330 | 0.5514 | -1.30 | 0.1704 | | |
| | 10 | -0.4902 | 0.5604 | -1.17 | 0.2109 | | |
| Single Mean | 0 | 0.6041 | 0.9759 | 0.60 | 0.9873 | 5.57 | 0.0312 |
| | 1 | -0.1836 | 0.9417 | -0.13 | 0.9366 | 2.90 | 0.3533 |
| | 2 | -0.1450 | 0.9418 | -0.10 | 0.9403 | 2.24 | 0.5143 |
| | 3 | 0.2464 | 0.9617 | 0.18 | 0.9663 | 1.90 | 0.5968 |
| | 4 | 0.7253 | 0.9784 | 0.65 | 0.9885 | 2.23 | 0.5173 |
| | 5 | 0.6332 | 0.9755 | 0.45 | 0.9812 | 1.21 | 0.7655 |
| | 6 | 0.8921 | 0.9822 | 0.63 | 0.9876 | 0.93 | 0.8362 |
| | 7 | 1.0907 | 0.9861 | 1.00 | 0.9950 | 1.39 | 0.7287 |
| | 8 | 0.8924 | 0.9816 | 0.85 | 0.9927 | 1.70 | 0.6577 |
| | 9 | 0.8269 | 0.9797 | 0.71 | 0.9894 | 1.26 | 0.7595 |
| | 10 | 1.0064 | 0.9835 | 1.14 | 0.9964 | 1.60 | 0.6799 |
| Trend | 0 | -4.0739 | 0.8684 | -1.41 | 0.8380 | 1.67 | 0.8441 |
| | 1 | -6.8854 | 0.6306 | -1.66 | 0.7431 | 1.49 | 0.8778 |
| | 2 | -8.3838 | 0.4936 | -1.67 | 0.7403 | 1.51 | 0.8732 |
| | 3 | -8.5802 | 0.4744 | -1.55 | 0.7882 | 1.48 | 0.8785 |
| | 4 | -6.0929 | 0.6997 | -1.29 | 0.8693 | 1.53 | 0.8700 |
| | 5 | -15.7044 | 0.0863 | -1.61 | 0.7597 | 1.91 | 0.7972 |
| | 6 | -241.768 | 0.0001 | -2.16 | 0.4894 | 3.55 | 0.5128 |
| | 7 | -331.888 | 0.0001 | -2.27 | 0.4348 | 4.63 | 0.3240 |
| | 8 | -25.1388 | 0.0019 | -1.51 | 0.7972 | 2.17 | 0.7531 |
| | 9 | 622.5544 | 0.9999 | -1.52 | 0.7888 | 1.99 | 0.7851 |
| | 10 | -67.2381 | <.0001 | -1.58 | 0.7658 | 3.04 | 0.6015 |

Figure 4.47: Results of the Augmented Dickey-Fuller (ADF) test of no differenced series

| Type | Lags | Rho | <u>Pr</u> < Rho | Tau | <u>Pr</u> < Tau | F | <u>Pr</u> > F |
|-------------|------|----------|-----------------|-------|-----------------|-------|---------------|
| Zero Mean | 0 | -45.2239 | <.0001 | -9.37 | <.0001 | | |
| Single Mean | 0 | -45.2373 | 0.0001 | -9.21 | 0.0002 | 42.43 | 0.0010 |
| Trend | 0 | -45.3074 | <.0001 | -9.05 | 0.0001 | 41.00 | 0.0010 |

Figure 4.48: Results of the Augmented Dickey-Fuller (ADF) test with first-order differencing

The series is found to be stationary by applying logarithm transformation and first-order differencing. Hence, impact analysis is carried out using the transfer function-noise model. At this stage, the main objective is to identify the transfer function representation and model the noise. The general representation of the transfer function model described

in Section 3.12 of Chapter 3 forms the basis. The time plot of the series shown in Figure 4.28 is examined further to decide the preliminary representation of the transfer function. This model is represented by the value of b , r and s respectively. Several types of transfer function are considered for the series based on the trend of the log rate of fatalities ever since the safety measure (intervention) is imposed. Three assumptions are considered for the model. Firstly, there is immediate or no delays between the response and the intervention applied. Secondly, the trend of the curve after the intervention follows an exponential ($\delta = 1$) or wave-like ($\delta = 2$) pattern. Finally, the value of s ranges from 0 to 1, which provides insight that the trend will begin to decay or increase from one or two years after the intervention.

$$\begin{aligned}
 (1) \quad v(B) &= \frac{\omega_0}{1 - \delta_1 B} I_t \quad b = 0, r = 1 \text{ and } s = 0 \\
 (2) \quad v(B) &= \frac{\omega_0 - \omega_1 B}{1 - \delta_1 B} I_t \quad b = 0, r = 1 \text{ and } s = 1 \\
 (3) \quad v(B) &= \frac{\omega_0}{1 - \delta_1 B - \delta_2 B^2} I_t \quad b = 0, r = 2 \text{ and } s = 0 \\
 (4) \quad v(B) &= \frac{\omega_0 - \omega_1 B}{1 - \delta_1 B - \delta_2 B^2} I_t \quad b = 0, r = 2 \text{ and } s = 1
 \end{aligned}$$

The cross-correlation pattern between the model's variables is also developed and analysed in order to determine the transfer function orders (b , r and s) and particularly, to determine if the orders have a value greater than the expected values mentioned above. The SAS programming syntax used for this purpose is shown in Figure 4.49 while the cross-correlation pattern of the variables is shown in Figure 4.50. It can be observed from Figure 4.50 that all the transfer function weights, particularly from lag 0 onwards, are insignificant (close to zero). Hence, the tentative values assumed for b and s are appropriate. In addition, the exponential pattern may be suitable for the decay of weights due to the fact that there are no remarkable weights since lag 0. In summary, these patterns

suggest that the preliminary order of $b = 0$, $r = 1$ and $s = 0$ or 1 , respectively. The values, however, should be adjusted accordingly if the model does not fulfil the adequacy checks.

Now that the tentative values for the orders of the transfer function model have been determined, the next step involves formulating the noise model (ARIMA). All of the possible forms of the noise model are considered to determine the appropriate models and selecting the best among them. In practice, the sample ACF and PCF plots are usually referred to determine the autoregressive (AR(p)) and moving average orders (MA(q)) of the model. In this study, however, to provide an extensive analysis as well as to attain a more accurate model, various p and q orders are included in the modelling. This is combined with various combinations of b , r and s of the transfer function. The objective here is to search for combinations of the transfer function and noise model which will give the lowest SBC.

Figure 4.51 shows the programming syntax used to determine the potential transfer function-noise model. Screening is carried out after the SAS program has been executed and outputs are generated. This includes tuning the parameters of the model in order to refine the residuals, ensuring that the residuals are ultimately white noise. The parsimony of the model needs to be compared with the minimum information criteria (e.g. AIC and SBC). In this study, the SBC value is used as the information criterion, as mentioned previously. The optimum model is the model with the lowest SBC.

```

data safety;
  input year logfatalities;
  opsikap = (year ge 2001);
  datalines;
1981    0.977
1982    1.001
1983    0.993
1984    0.965
1985    0.929
1986    0.897
1987    0.858
1988    0.843
1989    0.872
1990    0.870
.....
2005    0.622
;

proc arima data=safety;
  identify var=opsikap noprint;
  identify var=logfatalities(1) crosscorr=opsikap(1);
run;

```

Figure 4.49: SAS syntax used to determine the cross-correlation pattern of the model's variables

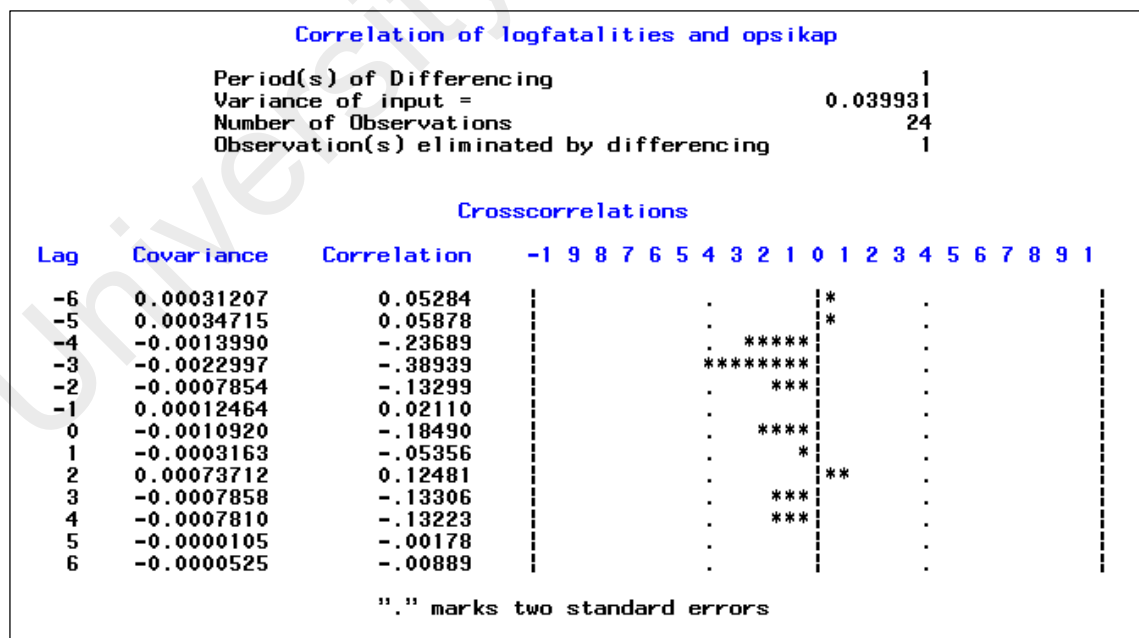


Figure 4.50: Cross-correlation pattern between the variables of the model

It can be seen from Figure 4.51 that the transfer function-noise model includes ‘opsikap’ as the intervention variable. This variable represents the Integrated Road Safety Operations (*Ops Sikap*) implemented since year 2001. For this model, it is considered that the variable ‘opsikap’ has influenced the series since year 2001 onwards. The data series analysed is the rate of road traffic log fatalities from 1981 to 2005. First-order differencing is applied to attain a white noise series, while maximum likelihood method is used to estimate the model. The order of the $AR(p)$ and $MA(q)$ is from 0 to 12, resulting in the generation of 169 different models by the software. Based on the analysis for both time series plots and cross-correlation checks, $b = 0$, $r = 1$ and $s = 0$. However, it is worth mentioning that sometimes the transfer function orders suggested by the plots and cross-correlation checks may not give the appropriate models. Hence, various orders, especially r (1, 2) and s (0, 1, 2), have been executed using the statistical software. Various sets of transfer function models are obtained using these combinations, whereby one set consists of 169 models.

In order to select the most appropriate among the 169 models produced, the first criterion is to assess the p -value of the parameters for each model. Preliminary screenings suggest that nine noise models (Table 4.7) require further checks since the parameters of these models have p -values less than 0.1. Hence, these nine potential transfer function-noise models are examined further in order to determine the best model. Moreover, the transfer function-noise (0,1,8) and (3,1,4) model outputs are chosen to demonstrate the steps taken to select the best model. This involves checking the estimate parameters, information criterion and performing autocorrelation check of the residuals for each potential model, as shown in Figure 4.52.

The next criterion used to determine whether a model is appropriate is white noise. Autocorrelation checks of the residuals reveal that the Q -statistic (chi-square statistic) is

greater than 0.05 for all of the nine tentative models. For instance, the Q -statistic of the transfer function-noise (0,1,8) and (3,1,4) models is greater than 0.05, as shown in Figure 4.52. This confirms that the transfer function-noise (0,1,8), (3,1,4) models are white noise. Thus, all of the nine models are tentatively adequate unless the adequacy checks deny this preliminary conclusion. Nevertheless, there is a need to sort the competing models based on their ability to forecast the series prior to carrying out the adequacy checks. The SBC of the models is used for this purpose. Table 4.8 shows the SBC value for each model, whereby the SBCs have been sorted from lowest to highest.

```

data safety;
  input year logfatalities;
  opsikap = (year ge 2001);
  datalines;
1981    0.977
1982    1.001
1983    0.993
1984    0.965
1985    0.929
1986    0.897
1987    0.858
1988    0.843
1989    0.872
1990    0.870
.....
2005    0.622
;

proc arima data=safety;
  i var=logfatalities(1) crosscorr=opsikap(1);
  /*--- (1) data first differencing nonseasonal d=1 ---*/
  e p=(0) q=(0) input=/(1) opsikap) method=ml;
  e p=(0) q=(1) input=/(1) opsikap) method=ml;
  e p=(0) q=(2) input=/(1) opsikap) method=ml;
  .....
  e p=(6) q=(10) input=/(1) opsikap) method=ml;
  e p=(6) q=(11) input=/(1) opsikap) method=ml;
  e p=(6) q=(12) input=/(1) opsikap) method=ml;
  .....
  e p=(12) q=(10) input=/(1) opsikap) method=ml;
  e p=(12) q=(11) input=/(1) opsikap) method=ml;
  e p=(12) q=(12) input=/(1) opsikap) method=ml;
run;

```

Figure 4.51: SAS programming syntax used to develop the potential transfer function-noise model

Based on the results shown Table 4.8, the transfer function-noise (12,1,4) model has the lowest SBC compared with other competing models. The model has an SBC of -94.077, while the SBC of the transfer function-noise (12,1,2) and (3,1,4) model is -93.527 and -91.760, respectively. Thus, based on the SBC value, the transfer function-noise (12,1,4) model is selected as the best model. However, the adequacy of the model needs to be checked before the model can be used for forecasting. The model with lowest SBC may not necessarily provide a stable forecast. Adequacy check is done by analysing the ACF and PACF for residuals and the cross-correlation between the residuals and the pre-whitened input series. The SAS syntax used for this purpose is shown in Figure 4.53, while some of the outputs are shown in Figure 4.54 and Figure 4.55.

Table 4.7: Transfer function-noise models with significant parameters

| No. | Transfer function orders | Model |
|-----|----------------------------|----------------------------------|
| 1 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (0,1,8) |
| 2 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (3,1,4) |
| 3 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (4,1,0) |
| 4 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (4,1,3) |
| 5 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (4,1,8) |
| 6 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (10,1,8) |
| 7 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (12,1,2) |
| 8 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (12,1,4) |
| 9 | $b = 0, r = 1$ and $s = 0$ | Transfer function-noise (12,1,5) |

Table 4.8: SBC of the transfer function-noise models

| No. | Model | SBC |
|-----|----------------------------------|---------|
| 1 | Transfer function-noise (12,1,4) | -94.077 |
| 2 | Transfer function-noise (12,1,2) | -93.527 |
| 3 | Transfer function-noise (3,1,4) | -91.760 |
| 4 | Transfer function-noise (12,1,5) | -91.347 |
| 5 | Transfer function-noise (4,1,3) | -91.321 |
| 6 | Transfer function-noise (4,1,0) | -89.414 |
| 7 | Transfer function-noise (4,1,8) | -88.099 |
| 8 | Transfer function-noise (0,1,8) | -87.806 |
| 9 | Transfer function-noise (10,1,8) | -85.805 |

Figure 4.54 shows a sample of the ACF, IACF and PACF of the residuals produced by the transfer function-noise (12,1,4) model. Checks on the ACF and PACF of the residuals confirm that the transfer function-noise (12,1,4) model is white noise. The shape of the ACF diminishes while the PACF is cut off after lag 1. Autocorrelation checks of the residuals also confirm that the model is white noise, as shown in Figure 4.55. The Q -statistic is greater than 0.05 at all lags. The adequacy check is continued by examining the cross-correlation checks of the model's residuals with 'opsikap' as the input variable. It can be observed from Figure 4.55 that the Q -statistic is also greater than 0.05 at all lags of the cross-correlation checks for the model's residuals. Consequently, the adequacy checks confirm that the model is adequate to represent the series and to be used for forecasting purposes.

| Maximum Likelihood Estimation | | | | | | | | | |
|------------------------------------|------------|----------------|------------|----------------------------|--------|---------------|--------|--------|--------|
| Parameter | Estimate | Standard Error | t Value | Approx Pr > t | Lag | Variable | Shift | | |
| MU | -0.01308 | 0.0031522 | -4.15 | <.0001 | 0 | logfatalities | 0 | | |
| MA1,1 | 0.99998 | 0 | Infty | <.0001 | 8 | logfatalities | 0 | | |
| NUM1 | -0.04946 | 0.02126 | -2.33 | 0.0200 | 0 | opsikap | 0 | | |
| DEN1,1 | -0.54711 | 0.27363 | -2.00 | 0.0456 | 1 | opsikap | 0 | | |
| Constant Estimate | | | | -0.01308 | | | | | |
| Variance Estimate | | | | 0.000565 | | | | | |
| Std Error Estimate | | | | 0.023761 | | | | | |
| AIC | | | | -92.3476 | | | | | |
| SBC | | | | -87.8056 | | | | | |
| Number of Residuals | | | | 23 | | | | | |
| Autocorrelation Check of Residuals | | | | | | | | | |
| To Lag | Chi-Square | DF | Pr > ChiSq | -----Autocorrelations----- | | | | | |
| 6 | 7.08 | 5 | 0.2151 | 0.095 | -0.038 | -0.168 | -0.429 | 0.047 | 0.102 |
| 12 | 11.28 | 11 | 0.4198 | 0.047 | -0.030 | -0.100 | -0.231 | 0.009 | 0.159 |
| 18 | 15.23 | 17 | 0.5788 | 0.131 | 0.063 | -0.124 | -0.136 | -0.023 | 0.017 |
| Maximum Likelihood Estimation | | | | | | | | | |
| Parameter | Estimate | Standard Error | t Value | Approx Pr > t | Lag | Variable | Shift | | |
| MU | -0.0098918 | 0.0022920 | -4.32 | <.0001 | 0 | logfatalities | 0 | | |
| MA1,1 | 0.81467 | 0.43607 | 1.87 | 0.0617 | 4 | logfatalities | 0 | | |
| AR1,1 | -0.35958 | 0.20851 | -1.72 | 0.0846 | 3 | logfatalities | 0 | | |
| NUM1 | -0.05887 | 0.02065 | -2.85 | 0.0044 | 0 | opsikap | 0 | | |
| DEN1,1 | 0.51976 | 0.29905 | 1.74 | 0.0822 | 1 | opsikap | 0 | | |
| Constant Estimate | | | | -0.01345 | | | | | |
| Variance Estimate | | | | 0.000576 | | | | | |
| Std Error Estimate | | | | 0.024003 | | | | | |
| AIC | | | | -97.437 | | | | | |
| SBC | | | | -91.7595 | | | | | |
| Number of Residuals | | | | 23 | | | | | |
| Autocorrelation Check of Residuals | | | | | | | | | |
| To Lag | Chi-Square | DF | Pr > ChiSq | -----Autocorrelations----- | | | | | |
| 6 | 2.21 | 4 | 0.6979 | 0.133 | 0.036 | -0.087 | 0.019 | -0.000 | -0.209 |
| 12 | 6.59 | 10 | 0.7634 | -0.183 | -0.013 | -0.036 | -0.249 | -0.077 | 0.044 |
| 18 | 7.20 | 16 | 0.9691 | 0.071 | 0.048 | -0.024 | -0.032 | 0.014 | 0.017 |

Figure 4.52: Estimated parameters, information criterion and autocorrelation check of residuals for the transfer function-noise (0,1,8) and (3,1,4) models

```

data safety;
  input year logfatalities;
  opsikap = (year ge 2001);
  datalines;
1981    0.977
1982    1.001
1983    0.993
.....
2005    0.622;

proc arima data=safety;
  identify var=opsikap(1) noprint;
  estimate p=(12) method=ml noprint;
  estimate q=(4) method=ml noprint; run;
  identify var=logfatalities(1) crosscor=(opsikap(1)) nlag=6; run;
  estimate p=(12) q=(4) input=/(1) opsikap method=ml; run;
  forecast out=fore
    id=year interval=1 printall; run;
  goptions cback=white colors=(black) border reset=(axis symbol) HTEXT=1.05;
  axis1 offset=(0.5 cm) label=('Year') minor=none
    order=(1996 to 2005);
  axis2 label=(angle=90 'Rate of Road Traffic Log Fatalities')
    order=(0.2 to 1.2 by 0.1);
  symbol1 i=join h=0.5 c=blue l=1 v=circle;
  symbol2 i=join h=1 c=red v=F;
  symbol3 i=join l=20;

proc gplot data=fore;
  plot logfatalities*year =1
    forecast*year        =2
    195*year              =3
    u95*year              =3 / overlay haxis=axis1 vaxis=axis2
                                vminor=0
                                href=2001
                                lh=2; run;

```

Figure 4.53: SAS syntax used to determine the ACF and PACF of the residuals, and the cross-correlation between the residuals and forecasts

Referring to the SAS outputs in Figure 4.56, the mathematical representation of the transfer function-noise (12,1,4) model can be formulated as in Equation (4.2).

$$Y_t = -0.00757 + \frac{(-0.07774)}{1 - 0.42057B} I_t + \frac{1 - 0.40226B^4}{(1 - 0.86064B^{12})(1 - B)} a_t \quad (4.2)$$

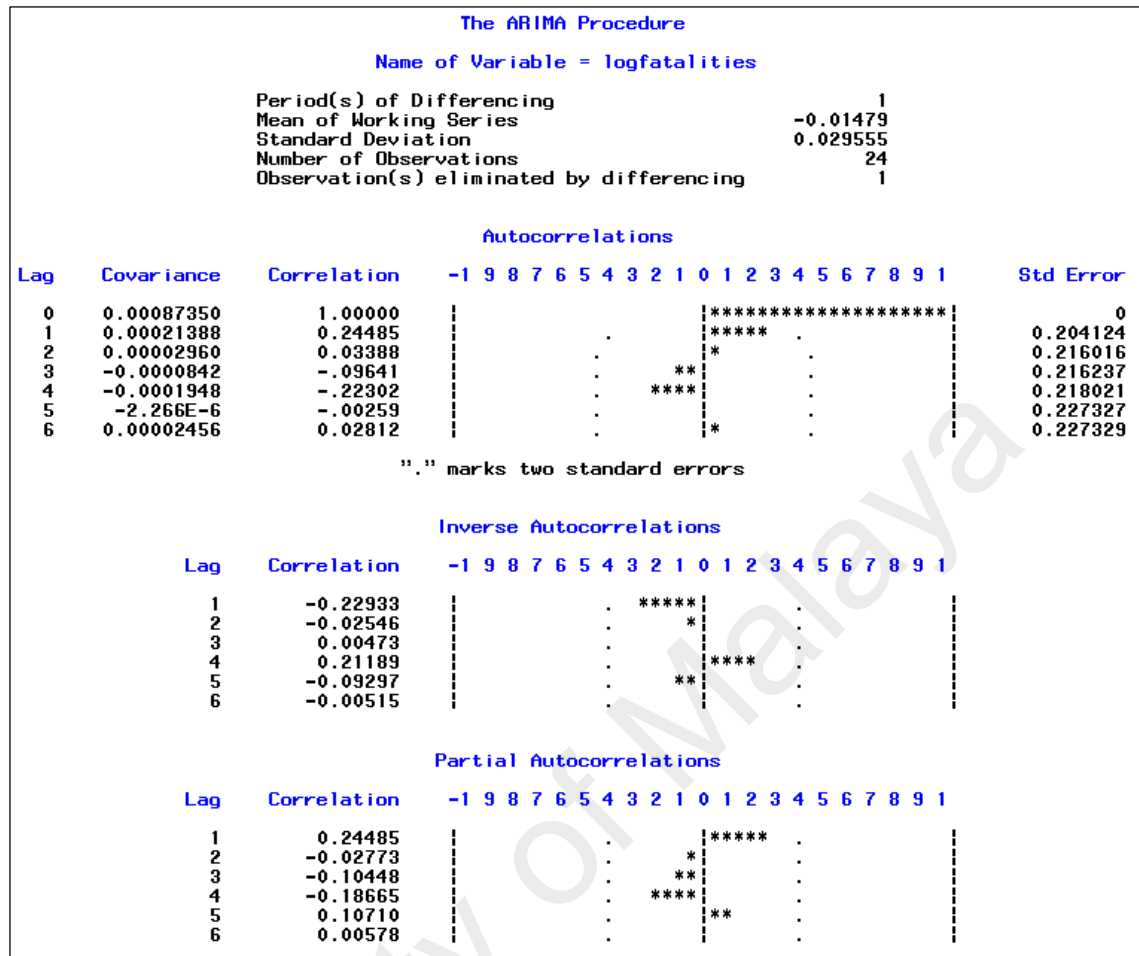


Figure 4.54: Sample of ACF and PACF residual plot for the transfer function-noise (12,1,4) model

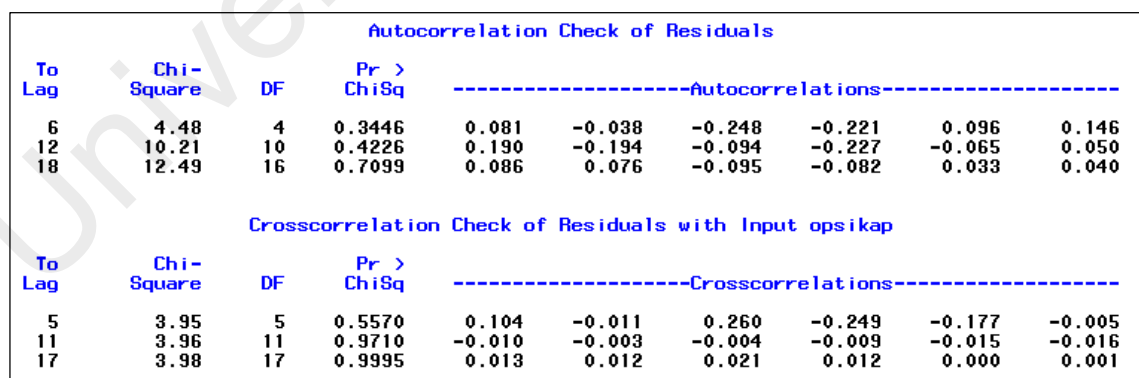


Figure 4.55: Autocorrelation check of residuals and cross-correlation check of residuals for the transfer function-noise (12,1,4) model

| The ARIMA Procedure | | | | | | | |
|----------------------------------|------------|----------------|---------|----------------|-----|---------------|-------|
| Maximum Likelihood Estimation | | | | | | | |
| Parameter | Estimate | Standard Error | t Value | Approx Pr > t | Lag | Variable | Shift |
| MU | -0.0075741 | 0.0036141 | -2.10 | 0.0361 | 0 | logfatalities | 0 |
| MA1,1 | 0.50302 | 0.20018 | 2.51 | 0.0120 | 4 | logfatalities | 0 |
| AR1,1 | 0.83400 | 0.13236 | 6.30 | <.0001 | 12 | logfatalities | 0 |
| NUM1 | -0.07782 | 0.01307 | -5.95 | <.0001 | 0 | opsikap | 0 |
| DEN1,1 | 0.36806 | 0.14595 | 2.52 | 0.0117 | 1 | opsikap | 0 |
| Constant Estimate | | | | -0.00126 | | | |
| Variance Estimate | | | | 0.000337 | | | |
| Std Error Estimate | | | | 0.018351 | | | |
| AIC | | | | -99.7546 | | | |
| SBC | | | | -94.0771 | | | |
| Number of Residuals | | | | 23 | | | |
| Model for variable logfatalities | | | | | | | |
| Estimated Intercept | | | | -0.00757 | | | |
| Period(s) of Differencing | | | | 1 | | | |
| Autoregressive Factors | | | | | | | |
| Factor 1: 1 - 0.834 B**(12) | | | | | | | |
| Moving Average Factors | | | | | | | |
| Factor 1: 1 - 0.50302 B**(4) | | | | | | | |
| Input Number 1 | | | | | | | |
| Input Variable | | | | opsikap | | | |
| Period(s) of Differencing | | | | 1 | | | |
| Overall Regression Factor | | | | -0.07782 | | | |
| Denominator Factors | | | | | | | |
| Factor 1: 1 - 0.36806 B**(1) | | | | | | | |

Figure 4.56: Parameters of the transfer function-noise (12,1,4) model

It can be seen from Figure 4.56 above that the constant (intercept) of the model is too small, i.e. -0.00757. However, the error of the constant is also relatively small (0.00361). Thus, the constant can be removed from the model to simplify the model without imposing a major effect on the model. Indeed, for this case, the removal of the constant provides a model with lower SBC, as shown in Figure 4.57.

| The ARIMA Procedure | | | | | | | |
|----------------------------------|----------|----------------|---------|----------------|----------|---------------|-------|
| Maximum Likelihood Estimation | | | | | | | |
| Parameter | Estimate | Standard Error | t Value | Approx Pr > t | Lag | Variable | Shift |
| MA1,1 | 0.40226 | 0.16701 | 2.41 | 0.0160 | 4 | logfatalities | 0 |
| AR1,1 | 0.86064 | 0.08722 | 9.87 | <.0001 | 12 | logfatalities | 0 |
| NUM1 | -0.07774 | 0.01240 | -6.27 | <.0001 | 0 | opsikap | 0 |
| DEN1,1 | 0.42057 | 0.12706 | 3.31 | 0.0009 | 1 | opsikap | 0 |
| Variance Estimate | | | | 0.000325 | | | |
| Std Error Estimate | | | | 0.018022 | | | |
| AIC | | | | -99.4295 | | | |
| SBC | | | | -94.8875 | | | |
| Number of Residuals | | | | 23 | | | |
| Model for variable logfatalities | | | | | | | |
| Period(s) of Differencing | | | | | 1 | | |
| No mean term in this model. | | | | | | | |
| Autoregressive Factors | | | | | | | |
| Factor 1: 1 - 0.86064 B**(12) | | | | | | | |
| Moving Average Factors | | | | | | | |
| Factor 1: 1 - 0.40226 B**(4) | | | | | | | |
| Input Number 1 | | | | | | | |
| Input Variable | | | | | opsikap | | |
| Period(s) of Differencing | | | | | 1 | | |
| Overall Regression Factor | | | | | -0.07774 | | |
| Denominator Factors | | | | | | | |
| Factor 1: 1 - 0.42057 B**(1) | | | | | | | |

Figure 4.57: Parameters of the transfer function-noise (12,1,4) model without the constant

The mathematical representation of the model after removing the constant is given by:

$$Y_t = \frac{(-0.07774)}{1 - 0.42057B} I_t + \frac{1 - 0.40226B^4}{(1 - 0.86064B^{12})(1 - B)} a_t \quad (4.3)$$

Furthermore, the initial decay of the log fatalities when the safety measure (*Ops Sikap*) is implemented is $-\omega_0$ while in the long term, the intervention shall reduce the log fatalities by about $\frac{\omega_0}{1 - \delta_1} = \frac{-0.07774}{1 - 0.42057} = -0.134$.

The transfer function-noise (12,1,4) model, as shown in Equation (4.3), is used to forecast the rate of road traffic log fatalities from 1981 to 2005. The forecasted rates of log fatalities are presented in Figure 4.58, while the time plots of the actual and forecasted rates are shown in Figure 4.59.

| Obs | Forecast | Std Error | 95% Confidence Limits | | Actual | Residual |
|-----|----------|-----------|-----------------------|--------|--------|----------|
| 3 | 1.0010 | 0.0381 | 0.9262 | 1.0758 | 0.9930 | -0.0080 |
| 4 | 0.9930 | 0.0381 | 0.9182 | 1.0678 | 0.9650 | -0.0280 |
| 5 | 0.9650 | 0.0381 | 0.8902 | 1.0398 | 0.9290 | -0.0360 |
| 6 | 0.9290 | 0.0381 | 0.8542 | 1.0038 | 0.8970 | -0.0320 |
| 7 | 0.8998 | 0.0358 | 0.8296 | 0.9699 | 0.8580 | -0.0418 |
| 8 | 0.8677 | 0.0358 | 0.7976 | 0.9378 | 0.8430 | -0.0247 |
| 9 | 0.8555 | 0.0358 | 0.7853 | 0.9256 | 0.8720 | 0.0165 |
| 10 | 0.8831 | 0.0358 | 0.8129 | 0.9532 | 0.8700 | -0.0131 |
| 11 | 0.8937 | 0.0315 | 0.8320 | 0.9554 | 0.8640 | -0.0297 |
| 12 | 0.8850 | 0.0315 | 0.8232 | 0.9467 | 0.8580 | -0.0270 |
| 13 | 0.8603 | 0.0315 | 0.7986 | 0.9220 | 0.8410 | -0.0193 |
| 14 | 0.8572 | 0.0315 | 0.7955 | 0.9189 | 0.9230 | 0.0658 |
| 15 | 0.9200 | 0.0190 | 0.8828 | 0.9572 | 0.9180 | -0.0020 |
| 16 | 0.8975 | 0.0190 | 0.8603 | 0.9346 | 0.9140 | 0.0165 |
| 17 | 0.8856 | 0.0190 | 0.8484 | 0.9228 | 0.8670 | -0.0186 |
| 18 | 0.8308 | 0.0190 | 0.7936 | 0.8680 | 0.7980 | -0.0328 |
| 19 | 0.7652 | 0.0182 | 0.7296 | 0.8008 | 0.7660 | 0.0008 |
| 20 | 0.7471 | 0.0182 | 0.7115 | 0.7827 | 0.7550 | 0.0079 |
| 21 | 0.7090 | 0.0182 | 0.6733 | 0.7446 | 0.7140 | 0.0050 |
| 22 | 0.6915 | 0.0182 | 0.6559 | 0.7271 | 0.6900 | -0.0015 |
| 23 | 0.6708 | 0.0180 | 0.6354 | 0.7061 | 0.6910 | 0.0202 |
| 24 | 0.6769 | 0.0180 | 0.6416 | 0.7123 | 0.6560 | -0.0209 |
| 25 | 0.6369 | 0.0180 | 0.6016 | 0.6723 | 0.6220 | -0.0149 |

Figure 4.58: Actual and forecasted log rate of fatalities produced by the transfer function-noise (12,1,4) model

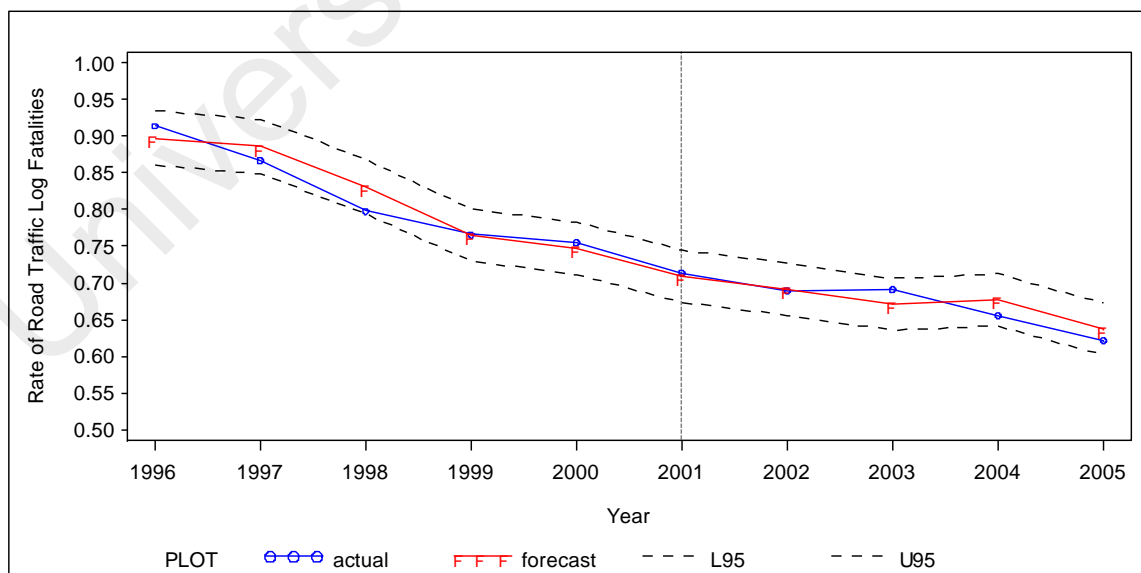


Figure 4.59: Time plot of actual and forecasted log rate of fatalities produced by the transfer function-noise (12,1,4) model

4.3.5 Impact of the road safety measures or interventions

Two types of impact owing to an intervention (road safety measure) can be attained from the transfer function-noise model: initial effect (ω_0) and long-term effect ($\omega_0/(1 - \delta_1)$ or $\omega_0/(1 - \delta_1 - \delta_2)$). Initially, both effects revealed by the software show the decay in the log rate of fatalities or accidents, but these effects are converted into the form of percentage by dividing each effect with the log rate of fatalities or accidents in a year before the intervention is implemented. For instance, the initial decay in the log rate of fatalities due to *Ops Sikap* conducted in 2001 is 0.07774, while its long-term effect is 0.134. The log rate of fatalities in year 2000 is 0.7471. Thus, the initial effect in the form of percentage is $0.07774/0.7471 \times 100 = 10.406\%$ whereas the long-term effect is $0.134/0.7471 \times 100 = 17.958\%$. In summary, there is a decline in the rate of fatalities of 10% due to the implementation of the Integrated Road Safety Operations (*Ops Sikap*) in 2001 in Malaysia. The initial impact is considered significant if the forecasted rate from the transfer function-noise model is outside the 95% confidence interval (CI) of the forecast by the ARIMA model.

The modelling is repeated for other measures in order to assess the initial and long-term effect of the road safety measures (interventions) implemented in Malaysia. The mathematical representations of the ARIMA and transfer function-noise models for each intervention are presented in Table 4.9, whereas the initial and long-term effects are shown in Table 4.10. The details of the parameters for both ARIMA and transfer function-noise are given in Appendix B. However, it can be seen from Table 4.10 that the long-term effect of some interventions is inapplicable, which is due to the fact the forecasted rate of fatalities for long-term periods is considered too low.

Table 4.9: Models used to assess the impact of road safety measures

| Intervention/Road safety measures | ARIMA model | Transfer function-noise model |
|--|--|---|
| Safety Helmet Rules in 1973 | The statistical software cannot be used for ARIMA modelling. The number of observations during pre-intervention is insufficient. | Y_t $= -0.01878$ $+ \frac{(0.14926 - 0.05875B + 0.0830)}{(1 + 0.0295B - 0.38981B)}$ $+ \frac{(1 + 0.11953B^5)}{(1 + 0.99997B^{12})(1 - B)} a_t$ |
| Effect of Seat Belt Rules in 1978 on car driver fatalities | ARIMA(8,1,0) $(1 - B)Y_t$ $= 0.03857$ $+ \frac{1}{(1 + 0.99999B^8)} a_t$ | Y_t $= \frac{(-0.08967) I_t}{(1 - 0.80031B)}$ $+ \frac{1}{(1 + 0.5955B)(1 - B)} a_t$ |
| Effect of Seat Belt Rules in 1978 on vulnerable road user fatalities | ARIMA(1,1,2) $(1 - B)Y_t$ $= 0.01084$ $+ \frac{(1 - 0.94964B^2)}{(1 + 0.99964B)} a_t$ | Y_t $= \frac{(-0.04702) I_t}{(1 - 0.92046B)}$ $+ \frac{(1 - 0.52648B^2)}{(1 + 0.92297B^1)(1 - B)} a_t$ |

Table 4.9, continued

| Intervention/Road safety measures | ARIMA model | Transfer function-noise model |
|---|---|--|
| Road Transport Act 1987 for Drunken Driving | <p>ARIMA(2,0,0)</p> $Y_t = -0.29137 + \frac{1}{(1 - 0.39749B^2)} a_t$ | $Y_t = -0.28936 + \frac{(-0.21060) I_t}{(1 - 1.11647B + 0.66378B^2)} + \frac{1}{(1 - 0.4389B^2)} a_t$ |
| Speed Limit Rules in 1989 | <p>ARIMA(11,0,8)</p> $Y_t = 2.42490 + \frac{(1 - 0.99927B^8)}{(1 + 0.42413B^{11})} a_t$ | $Y_t = 2.42774 + \frac{(-0.1336 + 0.15506) I_t}{(1 - 0.83678B)} + \frac{(1 - 0.77185B^8)}{(1 + 0.5271B^{11})} a_t$ |
| Motorcycle Daytime Running Headlight Regulation in 1992 | <p>ARIMA(0,2,1)</p> $(1 - B)^2 Y_t = (1 + 0.82635B) a_t$ | $Y_t = \frac{(0.10605) I_t}{1 - 0.69555B} + \frac{(1 + 0.84635B)}{(1 - B)^2} a_t$ |
| Specifications for Protective Helmets (MS 1:1996) | <p>Number of observations at pre-intervention is insufficient.</p> | $Y_t = 0.05409 + \frac{(-0.09054) I_t}{1 - 0.13052B - 0.75531B^2} + \frac{(1 + 0.9999B)}{(1 - 0.97801B^9)(1 - B)^2} a_t$ |

Table 4.9, continued

| Intervention/Road safety measures | ARIMA model | Transfer function-noise model |
|--|---|--|
| Road Safety Programmes in 1997 | $\text{ARIMA}(6,1,6)$ $(1 - B)Y_t$ $= \frac{(1 + 0.99505B^6)}{(1 + 0.99996B^6)} a_t$ | Y_t $= -0.00783$ $+ \frac{(-0.02595) I_t}{1 - 1.30133B + 0.90156B^2}$ $+ \frac{(1 - 0.63B^7)}{(1 + 0.99975B^{12})(1 - B)} a_t$ |
| Integrated Road Safety Operations (<i>Ops Sikap</i>) in 2001 | $\text{ARIMA}(12,1,4)$ $(1 - B)Y_t$ $= -0.00834$ $+ \frac{(1 - 0.51466B^4)}{(1 - 0.67104B^{12})} a_t$ | Y_t $= \frac{(-0.07774)}{1 - 0.42057B} I_t$ $+ \frac{1 - 0.40226B^4}{(1 - 0.86064B^{12})(1 - B)} a_t$ |
| Road Safety Plan 2006–2010 | $\text{ARIMA}(0,1,12)$ $(1 - B)Y_t$ $= (1 + 0.99899B^{12}) a_t$ | Y_t $= -0.01476$ $+ \frac{(-0.04538) I_t}{1 - 0.78322B + 0.61506B^2}$ $+ \frac{(1 + 0.99972B^{12})}{(1 - B)} a_t$ |

Table 4.10: Initial and long-term effect of the interventions on road safety

| Intervention/Road safety measures | Effect on the rate of road traffic fatalities | | |
|--|---|----------------|------------------|
| | Without safety measure | Initial effect | Long-term effect |
| Safety Helmet Rules in 1973 | NA | +43.114% | +78.375% |
| Effect of Seat Belt Rules in 1978 on car driver fatalities | +64.873% | -57.592%* | NA |
| Seat Belt Rules in 1978 linked with vulnerable road users | +9.123% | -5.807%* | -73.008% |
| Road Transport Act 1987 for Drunken Driving | -7.320% | -63.148% | NA |
| Speed Limit Rules in 1989 | +1.565% | -5.669% | +5.577% |
| Motorcycle Daytime Running Headlight Regulation in 1992 | -0.604% | +6.218% | +20.424% |
| Specifications for Protective Helmets (MS 1:1996) | NA | -30.816%* | NA |

Note: NA = not applicable, '+' = increase and '-' = decrease, *significant at 5% level

Table 4.10, continued

| Intervention/Road safety measures | Effect on the rate of road traffic injuries trend | | |
|--|---|----------------|------------------|
| | Without the safety measure | Initial effect | Long-term effect |
| Road Safety Programmes in 1997 | -0.526% | -5.675% | -9.454% |
| Integrated Road Safety Operations (<i>Ops Sikap</i>) in 2001 | +4.606% | -10.406%* | -17.958% |
| Road Safety Plan 2006–2010 | +3.274% | -7.090%* | -8.523% |

Note: NA = not applicable, '+' = increase and '-' = decrease, *significant at 5% level

4.4 Forecasting the number and rate of fatalities

A road safety plan is often complemented with a target on the number and/or rate of fatalities. Hence, a robust forecasting method is also developed along with the plan in order to provide the target. The forecasting model can also be used to check whether the programmes included in the plan are adequate to achieve the target. In this study, models are also developed to forecast the number and rate of fatalities. Number of fatalities from 1981 to 2012 are used to estimate the models. Whilst, number of fatalities in 2013 and 2014 are used to evaluate the forecasts. The objectives of the models are: (1) to attain the number and rate of fatalities beyond year 2020, (2) to determine when the target of 2.0 fatalities per 10,000 vehicles is achieved, (3) to determine whether the target of 5,358 deaths in 2020 is achievable, and (4) to provide a comparison with the forecasted fatalities by Rohayu et al. (2012). Both univariate and multivariate models are used to achieve these objectives. Both ARIMA and transfer function-noise models are classified as

univariate models, whereas multivariate analysis is carried out only for the state-space model.

4.4.1 ARIMA model

The steps taken for the ARIMA model are similar to the ARIMA modelling process mentioned above. However the primary difference is that the forecasted rate of fatalities is extended to year 2020. The ARIMA model is previously used to forecast the rate of fatalities or accidents until a year after the intervention has been implemented. The series of rate of fatalities from 1981 to 2012 is given in Appendix A. As a follow-up to the Box-Cox transformation check and ADF test, log transformation and first-order differencing are used to obtain a white noise (stationary) series. During the model estimation stage, ARIMA(2,1,0) is found to be one of the appropriate models and provides the lowest SBC. Hence, this model is used to forecast the rate of fatalities until year 2020. Equation (4.4) shows the mathematical representation of the ARIMA(2,1,0) model. The forecasted log rate of fatalities per 10,000 vehicles from 1981 to 2012 produced by the model is shown in Figure 4.60. It can be seen that the rate of fatalities predicted by the ARIMA model tends to have a steady decline beyond year 2012. The details of the model are given in Appendix B.

$$(1 - B)Y_t = \frac{1}{(1 - 0.29918B^2)} a_t \quad (4.4)$$

(0.090)

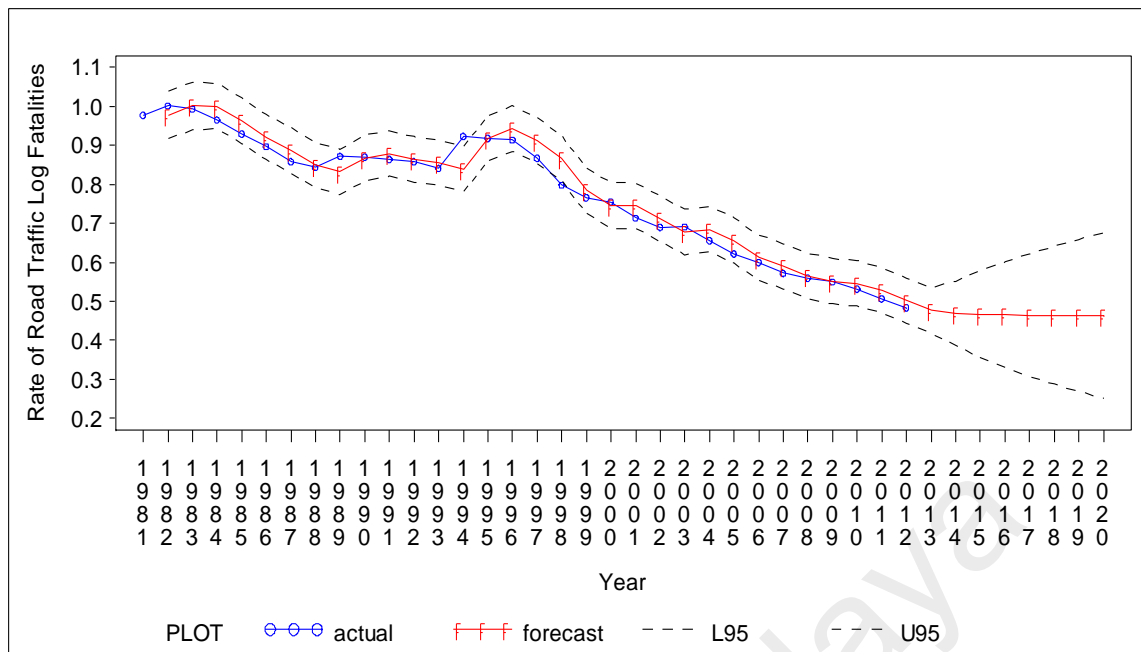


Figure 4.60: Actual and forecasted rate of fatalities from 1981 to 2020 produced by the ARIMA(2,1,0) model

4.4.2 Transfer function-noise model

Four circumstances are included in this modelling as the interventions that affect the log rate of fatalities from 1981 to 2012. These involve the increase in the rate of fatalities in 1994, road safety programmes in 1997, integrated road safety operations (*Ops Sikap*) since 2001 and finally the Road Safety Plan launched in 2006. The inclusion of the increase in fatalities rate in 1994 to the model becomes a factor that always increases the rate of fatalities. Basically, the rate of fatalities continues to increase since the exposures (population, vehicle registration, vehicle-kilometre travelled, road length, and etc.) increase constantly. However, the safety measures (interventions) result in a decrease in the rate of fatalities. In other words, the transfer function-noise model takes into account the fact that the rate of fatalities will increase in the absence of safety measures. However, it is notable that the level of decrease due to interventions is similar with that in the past. It is perceived that if the enforcement of safety measures in the future is different from that in past years, then the rate of fatalities will either increase or decrease. The results

reveal that the transfer function-noise (5,1,7) model is the best model to estimate the rate of fatalities until year 2020. The details of the model are shown in Table 4.11. Equation (4.5) shows the mathematical representation of the model, while the forecasted rates of log fatalities are depicted in Figure 4.61.

Table 4.11: Details of the transfer function-noise (5,1,7) model

| Parameter | Estimate | Standard Error | <i>t</i> -value | Approx. Pr > <i>t</i> | Lag | Variable |
|-----------|----------|----------------|-----------------|-------------------------|-----|---------------------|
| MU | -0.01764 | 0.00192 | -9.17 | <.0001 | 0 | Log fatalities rate |
| MA1,1 | -0.89732 | 0.30584 | -2.93 | 0.0033 | 7 | Log fatalities rate |
| AR1,1 | -0.93836 | 0.04944 | -18.98 | <.0001 | 5 | Log fatalities rate |
| NUM1 | 0.14011 | 0.00619 | 22.63 | <.0001 | 0 | Step 1994 |
| DEN1,1 | 0.29058 | 0.04562 | 6.37 | <.0001 | 1 | Step 1994 |
| NUM2 | -0.03031 | 0.00396 | -7.65 | <.0001 | 0 | Safety Prog. 1997 |
| DEN1,1 | 1.51190 | 0.07142 | 21.17 | <.0001 | 1 | Safety Prog. 1997 |
| DEN1,2 | -0.81182 | 0.06095 | -13.32 | <.0001 | 2 | Safety Prog. 1997 |
| NUM3 | -0.01784 | 0.00728 | -2.45 | 0.0142 | 0 | Ops. Sikap 2001 |
| DEN1,1 | 0.80296 | 0.11440 | 7.02 | <.0001 | 1 | Ops. Sikap 2001 |
| NUM4 | -0.01431 | 0.00680 | -2.10 | 0.0355 | 0 | Safety Plan 2006 |
| DEN1,1 | -0.96441 | 0.30626 | -3.15 | 0.0016 | 1 | Safety Plan 2006 |
| DEN1,2 | -0.72926 | 0.35630 | -2.05 | 0.0407 | 2 | Safety Plan 2006 |

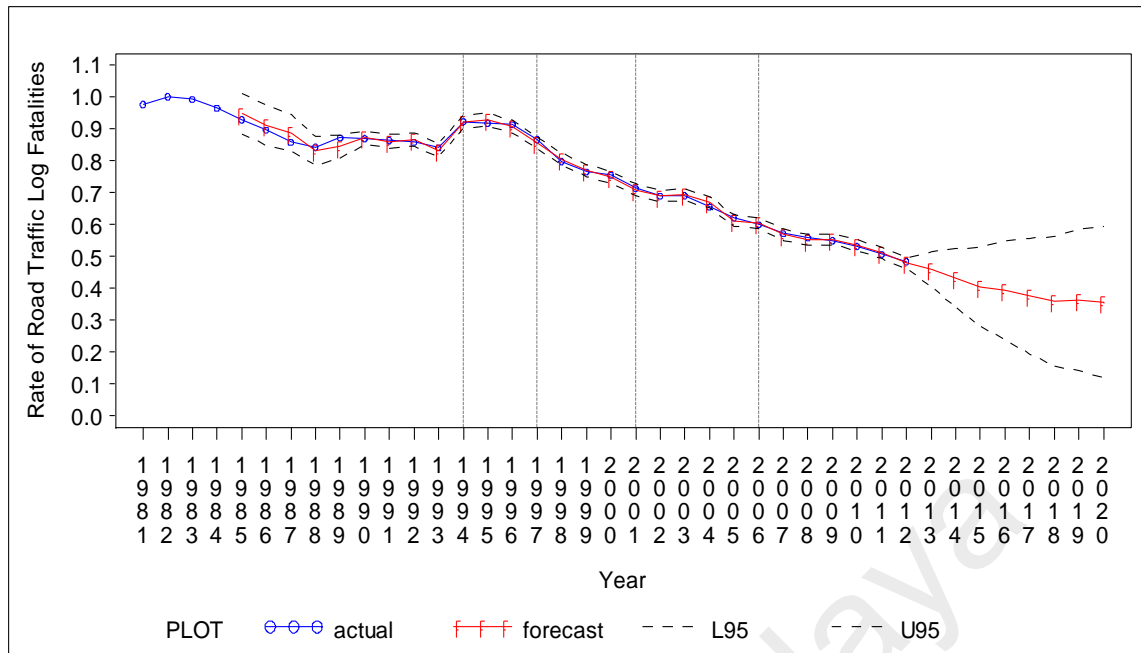


Figure 4.61: Actual and forecasted rate of fatalities from 1981 to 2020 produced by the transfer function-noise (5,1,7) model

$$\begin{aligned}
 Y_t = & -0.01764 + \frac{(0.14011) I_{1t}}{1 - 0.29058B} + \frac{(-0.03031) I_{2t}}{1 - 1.5119B + 0.81182B^2} + \frac{(-0.01784) I_{3t}}{1 - 0.80296B} \\
 & + \frac{(-0.01431) I_{4t}}{1 + 0.96441B + 0.72926B^2} + \frac{(1 + 0.89732B^7)}{(1 + 0.93836B^5)(1 - B)} a_t \quad (4.5)
 \end{aligned}$$

The forecasted rates of log fatalities produced by the transfer function-noise (5,1,7) model are compared with the rates obtained from the ARIMA(2,1,0) model. Figure 4.62 highlights this comparison. It can be seen that the transfer function-noise model predicts a lower rate of fatalities in the future compared with the prediction produced by the ARIMA(2,1,0) model. The transfer function-noise model forecasts that the log rate of fatalities is 0.357 in 2020, whereas the value is 0.464 according to the ARIMA(2,1,0) model.

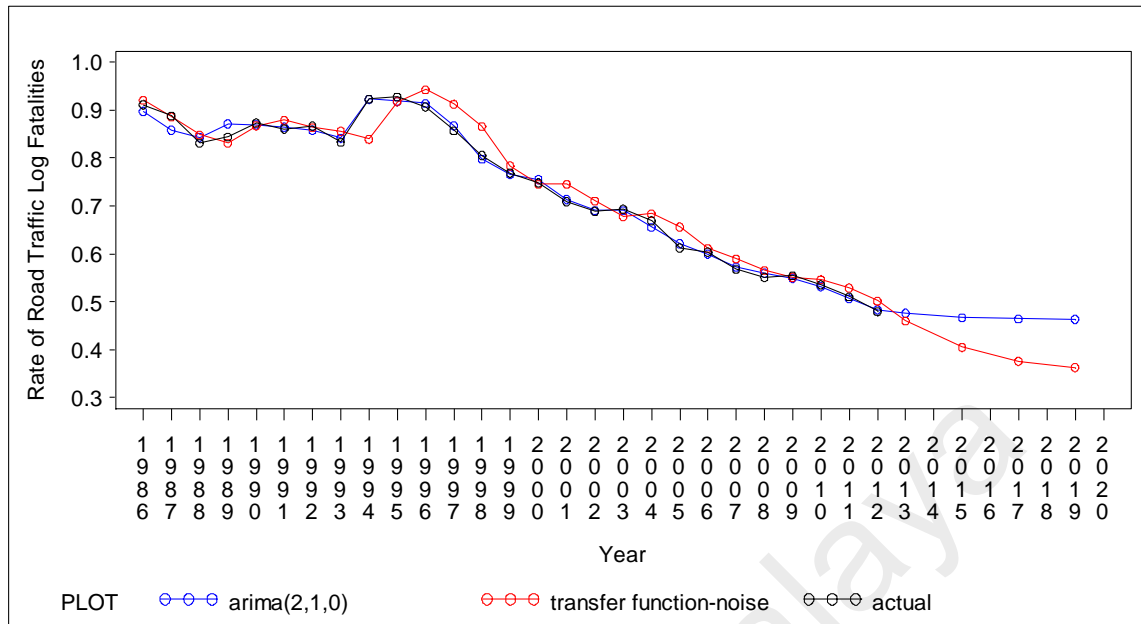


Figure 4.62: Actual and forecasted rate of fatalities up to year 2020 produced by the ARIMA(2,1,0) and transfer function-noise (5,1,7) models

In addition, if the log rate of fatalities in 2012 is 0.484, the reduction in the rate of fatalities from 2012 to 2020 is 7.6% according to the ARIMA(2,1,0) model. The transfer function-noise model, however, predicts a higher reduction (up to 25.5%) by year 2020. Nevertheless, it shall be noted that the transfer function-noise model adopts a level of enforcement which is similar and intense as the actual enforcement in past years.

4.4.3 Forecasting using the state-space model

A variety of explanatory variables that contribute significantly road traffic fatalities have been proposed in previous studies related to road safety. Some of these variables are presented in Chapter 2 of this thesis. Based on the comprehensive literature review given in Chapter 2, a total of 15 explanatory variables are expected to have a strong relationship with the number of fatalities in Malaysia. These variables are listed as follows.

- (1) GDP per capita (constant 2005 US\$)
- (2) Hospital beds (per 1,000 people)
- (3) Population (total, in 100,000)
- (4) Male population (% of total)
- (5) Female population (% of total)
- (6) Population age 15–24 years old (% of total)
- (7) Registered vehicles (in 10,000)
- (8) Registered motorcycles (in 10,000)
- (9) Registered buses (in 10,000)
- (10) Registered motorcycles (% of total)
- (11) Road sector energy consumption (kt of oil equivalent).
- (12) Road length (in 1,000 km).
- (13) Rural population (% of total)
- (14) Urban population (% of total)
- (15) Unemployment rate (in %)

In general, there is a large number of explanatory variables included in the model and therefore, it can be expected that there will be collinearity between these variables. For this reason, multi-collinearity tests are first carried out to assess the VIF values. Minitab statistical analysis software is used to determine the VIF. The results of the tests are presented in Table 4.12. Even though all of the 15 variables are included in the analysis, the Minitab software has removed two variables upon completion of the tests, namely female population and urban population. It is found that there is high correlation between these variables and other variables. The following alert is given by the Minitab software.

- * Female population (X5) is highly correlated with other X variables.
- * Female population (X5) has been removed from the equation.
- * Urban population (X15) is highly correlated with other X variables.
- * Urban population (X15) has been removed from the equation.

It can be seen from Table 4.12 that the VIF value of the variables varies, whereby most of the variables have a VIF greater than 10. Analysts recommend that the VIF should be less than 10 to prevent collinearity between the variables. Hence, it can be concluded from Table 4.12 that there is a strong relationship between the explanatory variables. The number of variables included in the model are reduced and this can be done by eliminating the variable(s) with the highest VIF value until the model has all variables with relatively small VIF (if possible, close to 1).

The stepwise method is used to simplify the process while retaining the quality of the model to be achieved at the same time. Statgraphics statistical software is used to run the regression analysis using the stepwise method. A p -value of 0.1 (10% confidence level) is adopted to judge whether a variable is adequate to be included in the model.

Table 4.12: VIF and other statistical parameters of explanatory variables

| Explanatory (predictor) variables | Coefficient | SE Coef. | <i>t</i> -statistic | <i>p</i> -value | VIF |
|-----------------------------------|-------------|----------|---------------------|-----------------|-----------|
| Constant | 41527 | 38591 | 1.08 | 0.298 | |
| GDP per capita (X1) | 1.7294 | 0.4871 | 3.55 | 0.003 | 344.138 |
| Hospital beds (X2) | 858 | 1220 | 0.70 | 0.492 | 52.675 |
| Population (X3) | 154.82 | 86.98 | 1.78 | 0.094 | 12482.719 |
| Male population (X4) | -2902 | 1115 | -2.60 | 0.019 | 626.922 |
| Population age 15-24 (X6) | 2681 | 1130 | 2.37 | 0.031 | 699.485 |
| Registered vehicles (X7) | -37.32 | 12.32 | -3.03 | 0.008 | 38430.553 |
| Registered motorcycles (X8) | 54.59 | 20.60 | 2.65 | 0.017 | 20886.837 |
| Registered buses (X9) | 2142.1 | 993.7 | 2.16 | 0.047 | 2477.602 |
| Registered motorcycles (X10) | -247.8 | 104.3 | -2.37 | 0.030 | 130.032 |
| Road energy consumption (X11) | 0.0330 | 0.1570 | 0.21 | 0.836 | 377.344 |
| Road length (X12) | 46.74 | 13.89 | 3.37 | 0.004 | 107.042 |
| Rural population (X13) | 639.8 | 381.6 | 1.68 | 0.113 | 11009.287 |
| Unemployment rates (X14) | -33.49 | 52.73 | -0.64 | 0.534 | 4.240 |

Table 4.13: Variables with p -value less than 0.1

| Explanatory variables | Estimate | Standard error | t -statistic | P -value |
|----------------------------------|----------|----------------|----------------|------------|
| Constant | 1235.88 | 2246.49 | 0.55014 | 0.5878 |
| GDP per capita (X1) | 1.04839 | 0.175467 | 5.97487 | 0.0000 |
| Hospital beds (X2) | 1035.53 | 492.443 | 2.10285 | 0.0471 |
| Registered vehicles (X7) | -3.39317 | 0.498503 | -6.80673 | 0.0000 |
| Registered buses (X9) | 777.378 | 215.474 | 3.60776 | 0.0016 |
| Registered motorcycles (%) (X10) | -60.123 | 30.7934 | -1.95246 | 0.0637 |
| Road length (X12) | 11.5389 | 4.74071 | 2.43399 | 0.0235 |

The results of the regression analysis suggest that six explanatory variables can be included in the model since they have a p -value less than 0.1, as shown in Table 4.13. Nevertheless, the VIF of the variables of the model with the six variables should be rechecked in order to avoid collinearity. The results are presented in Table 4.14 and it observed that most of the VIF values are still relatively high. Again, some variables are excluded from the model. Firstly, the number of registered buses is taken out due to the fact that it has the highest VIF (157.355). However, even though the number of registered buses variable has been removed from the model, the VIF still exceeds the recommended value. Hence, the number of registered vehicles is removed from the model. Again, the VIF value of the variables is relatively high. Thus, the GDP per capita is excluded from the model. Finally, model that includes the following variables (hospital beds, percentage of registered motorcycles and road length) has a VIF less than 10 (Table 4.15). Thus, these three variables are considered appropriate to be included in the state-space model

to predict the number of road traffic fatalities until 2020 using the multivariate analysis method. SAS statistical software is used for this purpose.

Table 4.14 VIF: and other statistical parameters of model with six variables

| Explanatory variables | Coefficient | SE Coef. | <i>t</i> -statistic | <i>p</i> -value | VIF |
|------------------------------|-------------|-------------|---------------------|-----------------|---------|
| Constant | 1798 | 2515 | 0.71 | 0.482 | |
| GDP per capita (X1) | 0.9915 | 0.1963 | 5.05 | 0.000 | 58.504 |
| Hospital beds (X2) | 543.4 | 526.1 | 1.03 | 0.312 | 9.212 |
| Registered vehicles (X7) | -3.3947 | 0.5693 | -5.96 | 0.000 | 91.052 |
| Registered buses (X9) | 793.1 | 244.9 | 3.24 | 0.003 | 157.355 |
| Registered motorcycles (X10) | -49.81 | 34.77 | -1.43 | 0.165 | 14.356 |
| Roads length (X12) | 12.492 | 5.300 | 2.36 | 0.027 | 18.716 |

Table 4.15: VIF and other statistical parameters of model with three variables

| Explanatory variables | Coefficient | SE Coef. | <i>t</i> - statistic | <i>p</i> -value | VIF |
|------------------------------|-------------|-------------|-------------------------|-----------------|-------|
| Constant | 17247 | 1908 | 9.04 | 0.000 | |
| Hospital beds (X2) | -2833.9 | 649.2 | -4.37 | 0.000 | 3.337 |
| Registered motorcycles (X10) | -132.57 | 44.50 | -2.98 | 0.006 | 5.593 |
| Roads length (X12) | 9.414 | 4.368 | 2.16 | 0.040 | 3.025 |

In the state-space modelling technique, the data series analysed must be stationary. Hence, the ADF unit root test is conducted for all three variables. After all the series are

white noise, the modelling can be carried out. The SAS statements used to complete the modelling process are rather long, and are given in Appendix C. Figure 4.63 shows the preliminary estimates for the state-space model developed, while the fitted model is shown in Figure 4.64.

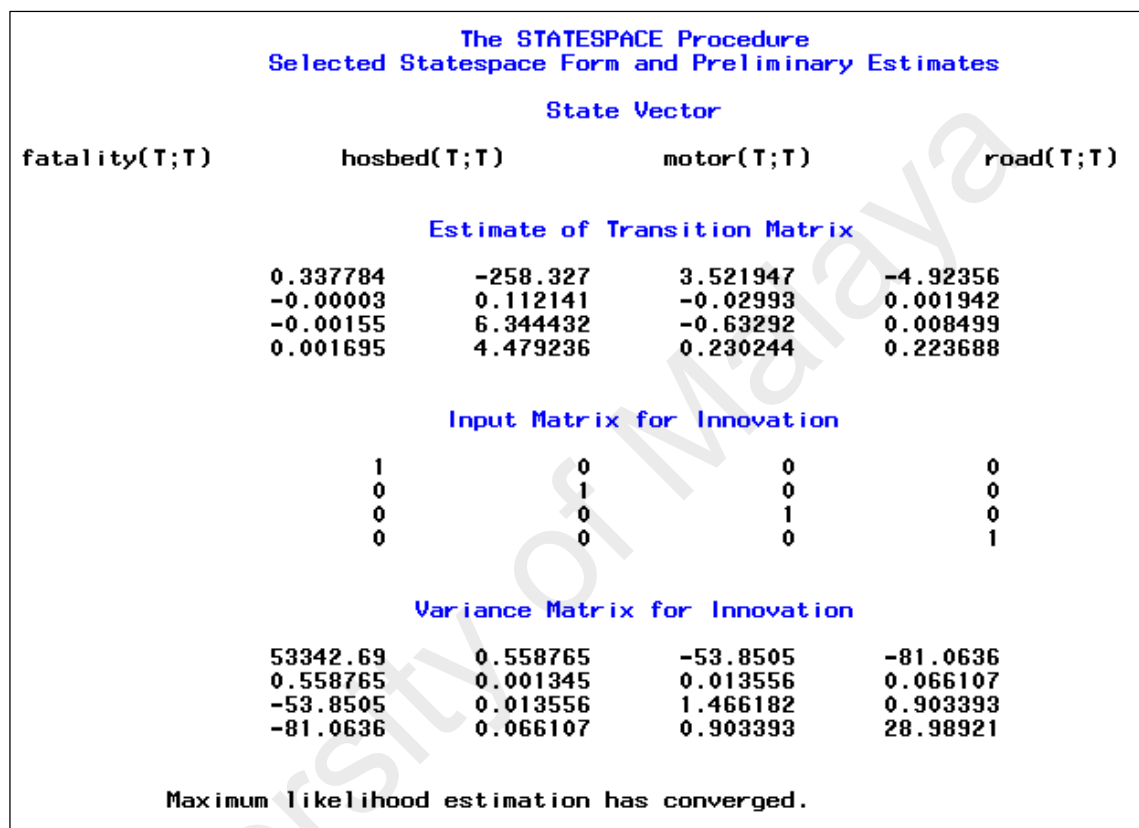


Figure 4.63: Preliminary estimates for the state-space model

Based on Figure 4.63, it is confirmed that the maximum likelihood of the state-space model reached convergence. Thus, the convergence problem is not encountered, the model is stable and can be used for forecasting purposes. Figure 4.65 shows the forecasted number of fatalities until year 2020 obtained from the state-space model. Since the other models use the rate of fatalities rather than the number of fatalities, the forecasted number of fatalities has been transformed into the rate of fatalities, and the results are shown in

Figure 4.65. Comparisons on the forecasted rate of fatalities revealed by the univariate and multivariate analysis described above are discussed further in the following chapter.

| The STATESPACE Procedure | | | |
|---|-------------|----------------|-----------|
| Selected Statespace Form and Fitted Model | | | |
| State Vector | | | |
| fatality(T;T) | hosbed(T;T) | motor(T;T) | road(T;T) |
| Estimate of Transition Matrix | | | |
| 0.337784 | -258.327 | 3.521947 | -4.92356 |
| -0.00003 | 0.112141 | -0.02993 | 0.001942 |
| -0.00155 | 6.344432 | -0.63292 | 0.008499 |
| 0.001695 | 4.479236 | 0.230244 | 0.223688 |
| Input Matrix for Innovation | | | |
| 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 1 |
| Variance Matrix for Innovation | | | |
| 53342.69 | 0.558765 | -53.8505 | -81.0636 |
| 0.558765 | 0.001345 | 0.013556 | 0.066107 |
| -53.8505 | 0.013556 | 1.466182 | 0.903393 |
| -81.0636 | 0.066107 | 0.903393 | 28.98921 |
| Parameter Estimates | | | |
| Parameter | Estimate | Standard Error | t Value |
| F(1,1) | 0.337784 | 0.167396 | 2.02 |
| F(1,2) | -258.327 | 5.679463 | -45.48 |
| F(1,3) | 3.521947 | 5.576596 | 0.63 |
| F(1,4) | -4.92356 | 4.519490 | -1.09 |
| F(2,1) | -0.00003 | 0.000028 | -1.27 |
| F(2,2) | 0.112141 | 0.137951 | 0.81 |
| F(2,3) | -0.02993 | 0.005502 | -5.44 |
| F(2,4) | 0.001942 | 0.001207 | 1.61 |
| F(3,1) | -0.00155 | 0.000904 | -1.72 |
| F(3,2) | 6.344432 | 3.666196 | 1.73 |
| F(3,3) | -0.63292 | 0.170279 | -3.72 |
| F(3,4) | 0.008499 | 0.039200 | 0.22 |
| F(4,1) | 0.001695 | 0.003990 | 0.42 |
| F(4,2) | 4.479236 | 5.495022 | 0.82 |
| F(4,3) | 0.230244 | 0.689915 | 0.33 |
| F(4,4) | 0.223688 | 0.174252 | 1.28 |

Figure 4.64: State-space model and fitted model

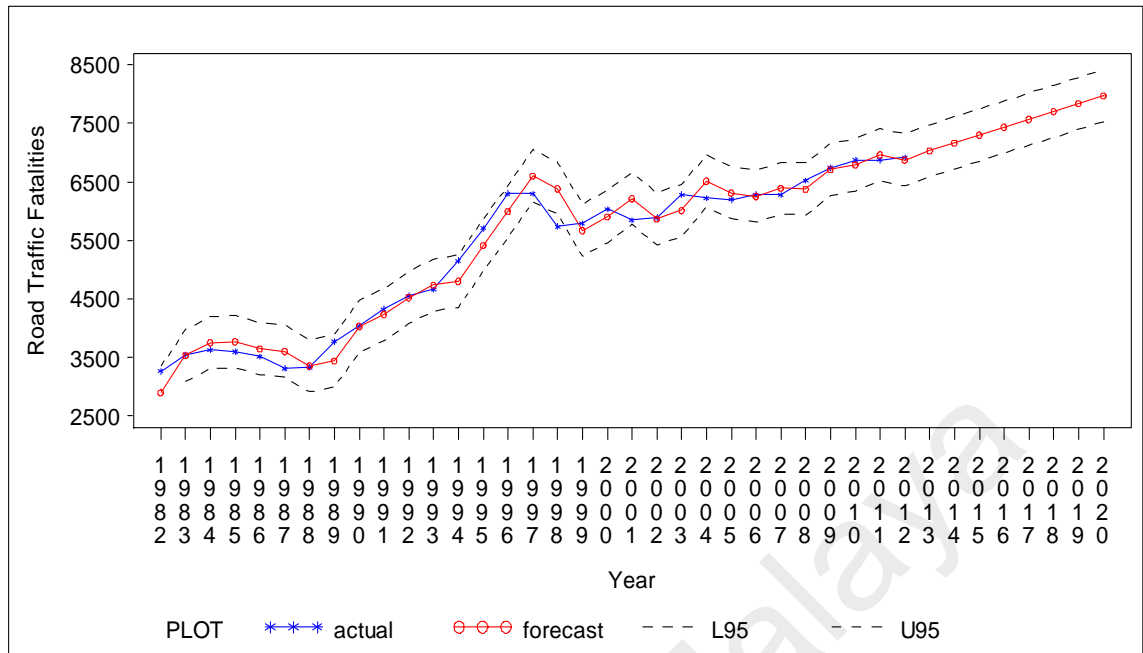


Figure 4.65: Actual and forecasted number of road traffic fatalities from 1982 to 2020 produced by the state-space model

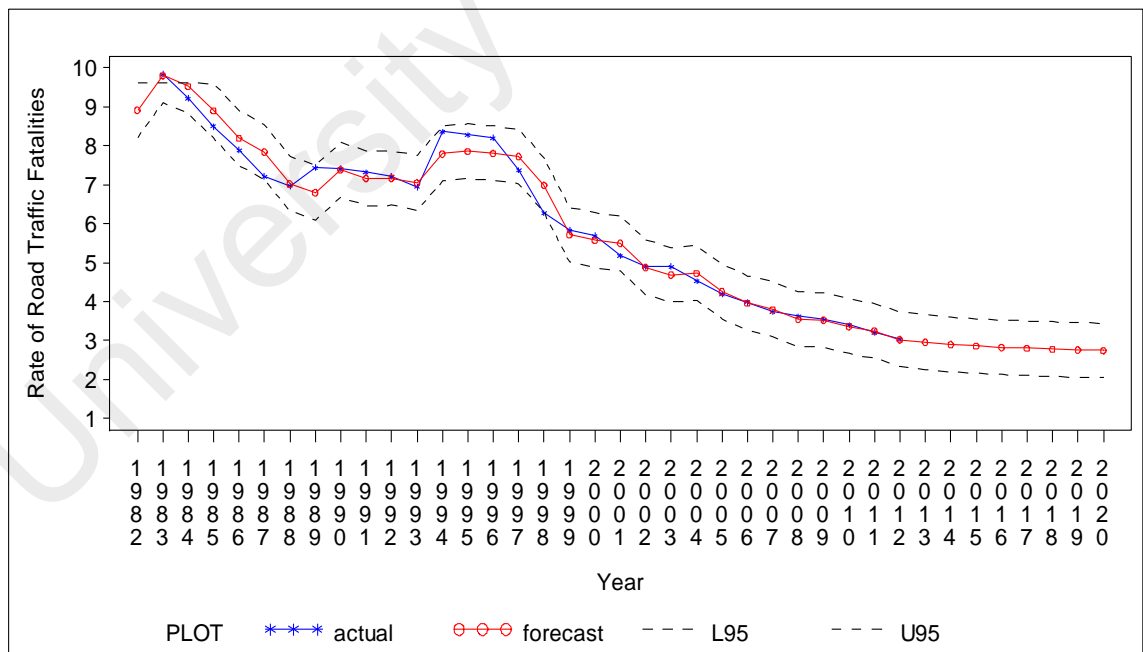


Figure 4.66: Actual and forecasted rate of road traffic fatalities from 1982 to 2020 produced by the state-space model

CHAPTER 5: DISCUSSION AND IMPLICATIONS OF THE STUDY

The results presented in the preceding chapter are discussed further in this chapter. The key findings are compared with those from other studies and the research questions raised in the first chapter are addressed. This chapter includes the forecasted fatalities in 2015 and 2020 produced by the models developed in this study.

5.1 Road traffic safety scenario in Malaysia

The key findings on the road traffic safety scenario in Malaysia are presented in this section. The headings of the following sub-sections are self-explanatory and they indicate the key findings of this study.

5.1.1 The major victims of road traffic fatalities are motorcyclists

In general, it is found that motorcar occupants are not the major victims of road traffic accidents – rather, motorcyclists are the most vulnerable road users and constitute the highest number of fatalities due to road traffic accidents. This trend is the same as that in other countries. For example, in Japan, pedestrians and cyclists constitute nearly 60% of the victims of road traffic accidents (Hayakawa et al., 2000). WHO (2004) reported that pedestrians, cyclists, and motorcyclists account for a large proportion of road traffic injuries in low-income and middle-income countries. Yan-Hong et al. (2006) discovered that pedestrians are the most common victims in Shanghai, with a percentage of 29.6%, followed by cyclists (25.1%) and motorcyclists (24.1%). In Denizli, Turkey, bicycles appear to be the means of transport that is most commonly involved in accidents, followed by (Sari et al., 2009). In Cambodia, the percentage of fatalities for motorcyclists is roughly 66% (Bachani et al., 2013).

5.1.2 The number and rate of fatal accidents are significantly higher in rural areas

In this study, it is found that the number of deaths due to road traffic accidents in rural areas is significantly higher than that in urban areas, which is consistent with the findings of previous studies (Clark & Cushing, 2004; Rakauskas et al., 2009; Travis et al., 2012). In contrast, Elvik et al. (2009) found that the accident rate in urban areas is 2 to 10 times higher than that in rural areas. The higher number of fatalities due to road traffic accidents in rural areas may be attributed to the decrease in the victims' chances for survival following severe injuries. In general, the probability of deaths and the probability of severe injuries increase with age, negligence on the use of safety belts, vehicle damage, high vehicle speeds, and early morning crashes. Travis et al. (2012) discovered that motorists with severe injuries are more likely to die in rural areas, after controlling for person-specific and event-specific factors.

Rakauskas et al. (2009) offered a plausible explanation for this scenario, which involves road design, proximity of emergency medical services and human factors. They also found that drivers in rural areas are more likely to engage in risky behaviour such as neglecting the use of seat belts since they have low perceptions on the risks associated with such a behaviour. In addition, rural drivers have low perception on the utility of government-sponsored traffic safety interventions compared to urban drivers. The factors mentioned above that have resulted in a higher number of fatalities in rural areas may also be relevant to the scenario in Malaysia. However, there are no comprehensive studies to date that have explored the contributory factors that lead to a higher number of road traffic deaths in rural areas in the country.

5.1.3 The percentage and rate of fatalities are higher for males

In most countries, the percentage of fatalities due to road traffic accidents is higher for males compared to females. This reflects that there is a possibility that males are exposed to higher risk factors of road traffic accidents. Stamatiadis and Deacon (1995) investigated the effect of gender on accident propensity and they discovered that on average, female drivers are safer compared to males. According to Abdel-Aty and Radwan (2000), male drivers have greater tendency to be involved in traffic accidents because of speeding. Dalvi (2004) also obtained a similar finding, whereby males are more likely to be involved in road traffic crashes compared to females. Mohammadi (2009) explored the usage of seat belts at two main roads in Iran. The results revealed that 52% male and 5% female drivers did not wear seat belts while driving, indicating that male drivers have higher negligence on road safety compared with females. Similarly, Sreedharan (2010) found that females are more likely to use helmets compared to males by a factor of 5.4. According to Bjørnskau (2011), male drivers have higher risk of being killed while driving compared to females. It is apparent from the findings of these studies that males are more likely to be involved in road traffic accidents which may be due to their tendency to drive at higher speeds and it can be expected that vehicle collisions at such speeds will be more deadly.

The higher number of fatalities for male drivers is also observed in studies in other countries. For instance, males constitute the majority of road traffic accident victims in Shanghai, with a percentage of over 69% (Yan-Hong et al., 2006). In China, the male : female ratio for fatalities due to road traffic crashes is 3.2 : 1 (Zhang et al., 2011). Kanchan et al. (2011) examined various epidemiological characteristics of road traffic accidents in South India and they found that 89.8% of the victims are males whereas the remaining 10.2% are females. In an analysis of the road safety in the US, Oster and Strong (2013)

concluded that the number of fatalities per 100,000 population is significantly lower for females compared to males for all age groups.

5.1.4 The highest percentage and rate of fatalities are those for young adult drivers/riders

Drivers/riders aged 16–25 years constitute the highest number of drivers/riders killed on roads, with a percentage of 35% of the total number of driver/rider fatalities in 2012. This is followed by drivers/riders aged 26–35 years and 36–45 years, with a percentage of 22 and 13.5%, respectively. This indicates that the highest number of victims of road traffic accidents in Malaysia are young adults. This finding is consistent with those from other countries. Ipingbemi (2008) found that more than 70% of the accident victims in Nigeria are those from the productive age group of 15–45 years. According to Evans (2004), the number of fatalities among drivers in the US reaches its peak for drivers aged 19 years old. Bener (2005) explored the trend of road traffic accidents in the State of Qatar and found that 43% of the total drivers who died are those without a valid licence within the age group of 10–19 years. In China, adults aged more than 65 years account for approximately 10% of the total road traffic deaths (Zhang et al., 2011).

In addition, the highest rates of fatalities are observed among young drivers within 16–20 years and the 21–24 years of age compared to other age groups in the US (Oster & Strong, 2013). It is believed that the risk curve according to age group follows a U-shaped trend which indicates that younger and older drivers are at higher risk of fatal accidents. Indeed, this is true since the findings of this study reveal that drivers/riders aged 16–25 years contribute to the highest percentage of fatalities. One possible explanation is that older drivers are less intense in using vehicles compared to young drivers and previous studies have indeed shown that young drivers within this age group are more at risk.

Zlatoper (1989) found that driving age has a statistically significant inverse effect on motor vehicle deaths and suggested that increasing the minimum driving age will help save lives. Stutts and Martell (1992) analysed data on motor vehicle crashes in North Carolina and observed that the crash rate declines significantly for drivers aged 55 and above compared to younger drivers.

In general, middle-aged drivers are safer than young and old drivers (Stamatiadis & Deacon, 1995). According to Cobb and Coughlin (1998), the most reckless drivers are those aged 16–24 years. Similarly, Kim et al. (1998) found that young and elderly drivers face up to three times the risk of being at fault compared to middle-aged drivers. According to Mayhew et al. (2006), numerous studies have confirmed that the risk of collisions as a function of age group follows a U-shaped relationship. The most common error made by senior drivers is the failure to yield the right-of-way, which arises from perceptual and attention problems, as well as misjudgement. Bossche et al. (2007) also confirmed that the risk curves are U-shaped, and they highlighted that male drivers aged 15–24 years old deserve special attention.

5.1.5 The highest percentage and rate of fatal accidents occurs at expressways

In this study, it is found that the highest rate of fatal accidents occurs at expressways, followed by federal roads and state roads. This trend is indeed expected since high vehicle speeds may be the main reason for the high rate of fatalities in expressways. In general, the maximum speed limit for expressways is 110 kph. In contrast, the maximum speed limit for federal roads and state roads is 90 kph. Relatively low speed limits are imposed for municipal roads and along designated areas such as schools and residential areas. The effect of higher vehicle speeds on the severity of accidents has been addressed in a number of studies such as Vernon et al. (2004) and Wilmot and Khanal (1999).

5.1.6 The highest percentage and rate of fatal accidents occurs at straight segments

It is found that about 75–80% of fatal accidents occur at straight segments in Malaysia. This may be due to the fact that the length of straight road segments is significantly higher compared to bends. Nevertheless, this finding can be used to alert the authorities on road safety that a large number of existing straight road segments are still inherently unsafe for road users. According to Jones et al. (2008), bends can cause crashes, but a recent study in the UK has shown that areas with many curved roads have lower crash rates compared to areas in which straight roads are dominant. The findings in the UK may be similar to the scenario in Malaysia since most of the road traffic deaths until year 2012 occurred on straight road segments rather than bend ones.

A large number of studies have shown that bend segments are more risky compared to straight ones. Milton and Mannering (1996) found that sharp horizontal curves have a potential effect on increasing the frequency of accidents. In another study, Milton and Mannering (1998) pointed out that younger drivers have greater tendency of being involved in accidents on roadway curves and especially during speeding. Horizontal curvature on its own does not increase the number of accidents – rather, it is dependent upon whether large straight sections precede the curves. Shankar et al. (1996) discovered that an increase in the number of horizontal curves per kilometre on rural freeways will increase the possibility of accidents.

Similarly, Kweon and Kockelman (2004) found that tighter horizontal curves increase fatal and injurious crash rates. In general, further analysis is required to conclude whether straight road segments are more hazardous than bend ones. For this reason, it is recommended that the rate of fatal accidents per length of bend segments is compared with the rate of accidents per length of straight segments. However, at present, even

though there are data available on the number of injuries at straight and bend road segments, there are no data on the length of bend and straight segments of the road networks in Malaysia. Hence, comparison cannot be made between the rate of accidents per length of bend segments and the rate of accidents per length of straight segments.

5.1.7 The highest percentage and rate of fatal accidents occur at T- and Y-junctions

It is found that the highest percentage of fatal accidents due to road traffic accidents occur at T- and Y-junctions regardless of the number of junctions in Malaysia. This is followed by the percentage of fatal accidents at cross (four-legged) junctions. In contrast, the least percentage of fatal accidents occur at interchanges and roundabouts. The high percentage of accidents at T- and Y- junctions may be attributed to the skew angles of the junctions, which restricts the visibility of motorists. According to Elvik et al. (2009), researchers have found that skewed junctions appear to be related to higher number of accidents compared to junctions with 90° angles between the roads. Skewed junctions are also most problematic for senior drivers. Furthermore, the accidents at junctions with 90° angles are less serious than those at skewed junctions. However, the finding in this study (i.e. the highest percentage of fatal accidents occurs at T- and Y- junctions) is contradictory to the findings of some studies. For example, Yannis et al. (2010) discovered that with respect to junction geometric design, 47% of fatal accidents occur at crossroads whereas 22% of fatal accidents occur at T- or Y- intersections.

In addition, a number of researchers have agreed that fewer road traffic accidents occur at grade-separated junctions and roundabouts (Elvik et al., 2009; O’Cinneide & Troutbeck, 1995). In contrast, O’Cinneide and Troutbeck (1995) found that T-junctions are safer compared to staggered junctions, but staggered junctions are generally safer

compared to cross junctions. This finding, however, contradicts with the current road safety scenario in Malaysia where staggered junctions are safer than T- and Y- junctions. It shall be noted that the data on the percentage of road traffic accidents at T- and Y- junctions are combined together in police reports. For this reason, it is not possible to make comparisons between the percentage of fatal accidents at staggered junctions and T-junctions in this study.

5.1.8 The percentage and rate of driver/rider fatalities are highest for single carriageways

It is found that more than 70% of driver/rider fatalities due to road traffic accidents occur on single carriageways from 1993 to 2012. However, it shall be noted that the comparisons made in this study may not be that accurate since the length for each type of road is not available. It may be possible that the percentage of driver/rider fatalities is highest on single carriageway roads due to the fact that this type of road is the longest, which leads to a higher frequency of accidents. To date, the results of previous studies pertaining to the number of lanes of a road segment is rather inconclusive. Milton and Mannering (1998) examined various geometric design elements and found that increasing the number of lanes on a given road segment leads to a higher number of accidents.

Noland and Oh (2004) also obtained similar results, such that an increase in the number of lanes and lane widths is associated with an increase in the number of fatalities. However, Kweon and Kockelman (2004) observed that fewer lanes result in higher fatal and injurious crash rates. In light of this discussion, the findings of Council and Stewart (1999) may shed some light regarding the number of lanes of a road and its consequence on the number of road traffic accidents. They examined the safety effects of converting two-lane rural roads into four-lane divided roads or four-lane undivided roads. The results

showed that there is a significant reduction in the number of accidents after the two-lane rural roads are converted into four-lane divided roads. However, the effect of converting two-lane rural roads into four-lane undivided roads on decreasing the number of accidents is not as pronounced. Hence, they proposed that separating lanes (or installing medians) may be relatively effective to reduce the number of accidents whereas merely increasing the number of lanes will not.

In addition, according to a comprehensive review by Elvik et al. (2009), there are no straightforward conclusions which can be drawn regarding the relationship between the number of lanes and the frequency of accidents since there is a large number of potential moderator variables which will influence the results. On the whole, the number of accidents may increase, but the severity of the accidents seems to decrease.

5.1.9 The probability of fatal accidents is higher during night time

In this study, it is found that the ratio of the percentage of fatal accidents between daytime and night time in 2011 is 54% : 46%. In general, the traffic volume in Malaysia is higher during daytime compared with night time. According to the traffic counts conducted at the Census Station BR501 in the suburbs of Selangor in 2011, the traffic volume during daytime is 58% whereas the traffic volume at night time is 42%. If the proportion of day and night traffic volumes at Census Station BR501 is representative of the whole country, it can be concluded that the probability of road traffic accidents is higher at night time. Nevertheless, further study is required to determine the probability of road traffic accidents during daytime and night time.

Neustrom and Norton (1993) forecasted that there is a higher number of fatal and injurious accidents at night time due to the increased likelihood of drunk driving. According to Elvik et al. (2009), the risk of accidents for motor vehicles is roughly 1.5–

2.0 times higher at night time compared to daytime. Around 35% of all injury accidents recorded in police reports in Norway occur during twilight and night time. The percentage is the same for both within and outside densely populated areas. However, the percentage of vehicles travelling along roads is only around 20–25% at night time. The risk of severe road traffic accidents is higher for those driving in the dark. In the US, the percentage of vehicles travelling along roads is roughly 25% at night time, and more importantly, the corresponding percentage of fatal accidents is 50%.

5.2 Reflection to the research questions

Based on the results presented in the previous chapter, it is now possible to answer the research questions presented in Chapter 1 regarding the road safety scenario in Malaysia.

5.2.1 How effective are the road safety measures implemented in Malaysia in reducing the percentage of casualties?

It is known that various road safety measures have been implemented in Malaysia such as legislations, standards, guidelines and safety programmes, as stipulated in the five-year road safety plan. A systematic analysis is required to assess the impact of the implementation of a particular road safety measure on the rate of road traffic casualties. Hence, univariate time series modelling is used for this purpose. The primary aim of a safety measure conducted by the government is to reduce the rate of road traffic casualties – however, the results in this study indicate that there is an increase in the rate of casualties after a safety measure has been implemented. In other words, the intervention may actually lead to an increase in the rate of casualties. Moreover, it is not the rate but the number of casualties that shall be reduced. Previously, the government used the “rate of

fatalities” as the target in Road Safety Plan, but recently it is replaced with “number of fatalities” following WHO target.

It shall be noted that the parameters used to assess the impact of a road safety measure in this study is not the same such as the rate of fatalities, the percentage of fatal accidents and the number of road traffic accidents. This is due to the availability of data and therefore, it is not possible to assess the impact of each road safety measure using the same parameters. For example, the statistics on road safety at the beginning of the data collection period are available only for Peninsular Malaysia. However, statistics on road safety in recent years are provided for both Peninsular Malaysia and East Malaysia. The number of observations at pre-intervention and post-intervention for certain data series are insufficient and therefore, the modelling process cannot be executed by the statistical software.

In addition, the impact of increasing the vehicle speed limits on the number road traffic accidents is also inconclusive. Researchers as mentioned in Chapter 2 perceive that the increase in vehicle speed limits will only increase the severity of the road traffic accidents and not the number of accidents. However, other researchers disagree with this point of view since they perceive that the increase in vehicle speed limits will also increase the number of accidents. In this study, only the impact of increasing vehicle speed limits on the number of accidents can be assessed due to the availability of data. Below are the significance of the road safety measures investigated in this study.

5.2.1.1 Safety Helmet Rules in 1973

The enactment of the Safety Helmet Rules in 1973 has influenced road safety in Malaysia at both individual and aggregate levels. According to the NHTSA report prepared by Wilson (1989), it has been estimated that the use of safety helmets is effective in preventing road traffic fatalities for motorcyclists by approximately 27%. Evans and Frick (1988) estimated that the use of safety helmets is effective to reduce fatalities by 28%. Liu et al. (2008) estimated that the use of safety helmets reduce the risk of deaths due to road traffic accidents by 42% (Liu et al., 2008). However, it shall be noted that the analysis conducted in this study is to examine the impact of safety helmet legislation on the national motorcycle fatalities at the aggregate level.

According to Watson et al. (1981), the enforcement of safety helmet laws results in a 30% decrease in the rate of fatalities whereas the repeal of such laws results in a 38% increase in the rate of fatalities. Graham and Lee (1986) provided estimates on the effectiveness of motorcycle helmet laws in reducing fatalities based on pooled time series and cross-sectional data in the US from 1975 to 1984. Estimates from the model suggest that the establishment of a helmet-wearing law decreases motorcycle mortality problems within a range of 12 to 22%. There is also evidence of gradual 'risk-compensation' behaviour. However, they did not include a survey on the percentage of motorcyclists wearing crash helmets in the NHTSA report. The percentage of helmet usage varies from 51 to 97 within this period.

Fleming and Becker (1992) used the time series intervention technique to analyse the impact of the Texas 1989 Helmet Law. The results showed that there is an initial decrease in the number of motorcyclists killed due to road traffic accidents, with a value of 12.6%. However, it shall be highlighted that there is also a 9% decline in the number of registered motorcycles within a year after the law was implemented. In 2000, the State of Florida

exempted adult motorcyclists and moped riders from wearing helmets provided that they have a medical insurance of \$10,000. Muller (2004) examined the impact of this law repeal and found that there is an increase in the rate of motorcycle occupant fatalities per registered motorcycle by 21.3% during the year when the law was amended. According to Al Haji (2005), many countries have legislated mandatory helmet use, which has been effective in preventing or reducing the severity of two-wheeler riders (motorcyclists and cyclists). The use of helmets varies from one country to another, and the usage rate tends to be high in high-income countries.

Mayrose (2008) suggested that if all states were to enact a primary motorcycle helmet law, there will be a dramatic increase in helmet use as well as a decrease in the number of motorcyclist head injuries and fatalities. According to Dee (2009), state laws which require helmet use appear to reduce motorcyclist fatalities by 27%. However, according to Elvik et al. (2009), introducing mandatory wearing of helmets reduces the number of injured moped riders and motorcyclists by around 26%. Hence, it is clear that the enactment of laws which make helmet use compulsory reduces the number of motorcycle casualties effectively.

In Malaysia, the use of helmets was made compulsory for motorcyclists on 1st March 1973. In this study, the impact of implementing this safety measure is analysed by modelling the rate of motorcyclist fatalities per 10,000 registered vehicles. However, it is found that the rate of motorcyclist fatalities actually increases after the enactment of the helmet law. In contrast, the rate of motorcyclist fatalities tends to decrease before this law is passed. The transfer function-noise model predicts that the rate of motorcyclist fatalities will increase by 43% ($p < 0.0001$) in 1973 (initial effect). However, the rate of motorcyclist fatalities will increase by 78% from 1973 onwards (long-term effect). It is

known that the percentage of registered motorcycles within 1972–1973 increases by 16% – however, there are no data available concerning the percentage of helmet usage.

Based on the results, it can be inferred that the Safety Helmet Rules imposed in 1973 will not effectively reduce the rate of motorcyclist fatalities, considering that there is an increase in the rate of motorcyclist fatalities after the helmet rules came into force. This finding contradicts the perception that helmet law is an effective public health policy. Xuequn et al. (2011) discovered that there is no significant increase in helmet use in the absence of such laws as well as public awareness. In Malaysia, the enforcement of the national safety helmet law in 2010 is at a level of 5 based on a scale of 0–10 (WHO, 2013). The enforcement of motorcycle helmet laws should be effectively supported by motorcycle safety programmes since one life per 1,000,000 inhabitants can be saved per year for every 10% increase in helmet usage (Abbas et al., 2012).

Four possible explanations can be linked with the increase in the rate of fatalities after the enactment of the helmet law. Firstly, even though it has been declared compulsory for motorcyclists to wear helmets since 1973, there is no significant increase in the percentage of helmet usage. This is specifically the case in rural areas in which the understanding on the benefits of safety helmet use and its level of enforcement may not be as intensive as in urban areas. In general, it is known that information, campaigns and enforcement on road safety are often conducted extensively in urban areas. According to Supramanian et al. (1984), earlier works on motorcycle safety research in Malaysia have shown that the use of crash helmets increases drastically following the enactment of helmet rules. However, Radin Umar et al. (1998) stated that only about 41% of motorcyclists use helmets – 25 years after the enactment of the helmet rules. The percentage of helmet use is saturated at 66% (Kamarudin et al., 2010).

Abbas et al. (2012) evaluated the effect of helmet use on the death rate of motorcycle riders on a global scale using data from 70 countries. They used a simple linear regression model to determine the correlation between helmet usage and rate of road traffic deaths, and the results showed that one life can be saved per 1,000,000 inhabitants each year for every 10% increase in helmet usage. The percentage of motorcyclists who refuse to wear crash helmets is the most significant factor which affects the death rate of motorcyclists. Thus, it can be inferred that the percentage of motorcyclists who wear crash helmets is very low at the time. The percentage of helmet usage among motorcyclists also increases gradually over the years.

Secondly, it is highly likely that motorcyclists tend to ride recklessly since they feel inherently safe after wearing crash helmets. This in turn, leads to an increase in motorcycle injuries. According to Garbacz (1989), there may be offsetting effects once a helmet law for motorcycle riders is enacted. The death toll of motorcyclists may increase because they tend to ride even more recklessly. The results of Ouellet (2011) do not fully support the risk compensation theory in motorcycle accidents based on data obtained from Los Angeles and Thailand. Their data refute the hypothesis that the increase in safety provided by crash helmets is offset by higher risk-taking while riding. The only evidence of risk compensation is that helmet use increases with greater amounts of travel. Thirdly, the quality of helmets at the time may not provide adequate protection for motorcyclists. Finally, the manner in which motorcyclists wear their helmets is also a probable cause. In 1998, only 54.4% of motorcyclists wore their safety helmets secured properly (Kulanthayan et al., 2000). It has been shown that the correct use of motorcycle helmets can reduce the risk of fatalities and risk of severe injuries by 40 and 70%, respectively (Lateef, 2011).

5.2.1.2 Seat Belt Rules in 1978

The first jurisdiction that mandates seat belt use is the State of Victoria in Australia in 1970 (Routley et al., 2009). The impact of the seat belt law has been greatly debated in the literature since then. Even though many researchers believe that the enactment of laws which make the use of seat belts compulsory will reduce road traffic fatalities, there are others who oppose this point of view. This is because they perceived that the use of seat belts leads to reckless driving or offsetting behaviour (Peltzman, 1975; Adams, 1982; Garbacz, 1990; McCarthy, 1999). Jaeger and Lassarre (2000) have also shown that there is a risk compensation effect, since the use of seat belts leads to an increase in speeding. Conversely, Freeman (1985) found that there is a reduction in fatal and serious injuries by 25% after the use of seat belts was made compulsory in the UK in 1983. Harvey and Durbin (1986) used structural time series analysis to explore the effect of compulsory use of seat belts on the safety of car occupants. They found that there is a reduction of 18% for car drivers and 25% for front seat passengers in terms of the number of road traffic fatalities.

Campbell et al. (1986) used the ARIMA modelling technique and discovered that the enforcement of the seat belt legislation in the US reduces the number of casualties – however, at a lower value than expected. This is possibly due to the level of seat belt use, which is within a range of 40–49% in most of the states in the US at the time. Wagenaar (1988) examined the effects of compulsory seat belt use on the number of occupants fatally injured in road traffic crashes. The results showed that there is a decrease in the rate of front-seat fatalities by 8.7%. He also observed that various studies on compulsory seat belt use leads to a decrease in fatalities up to 80%. With improved legislation and stricter law enforcement, NHTSA (1996) has shown that seat belts are effective in reducing fatalities by 45%.

In addition, Scuffham and Langley (2002) found that the 1984 universal seat belt wearing law results in a sustainable 15.6% reduction in fatal crashes. Mandatory seat belt use provides strong protection against fatalities in road traffic accidents in different countries according to various studies (Al Haji, 2005). Hermans et al. (2006) investigated the monthly frequency and severity of road traffic accidents in Belgium from 1974 to 1999. The introduction of the law of mandatory seat belt use for front seats yields a considerable increase in road safety. This law reduces all kinds of accidents and casualties. Analysing provincial data in Canada between 1980 and 1996, Sen and Mizzen (2007) suggested that roughly 17% of the observed decline in vehicle-occupant fatalities is attributable to the enactment of mandatory seat belt legislation and the corresponding increase in seat belt use.

Using quasi-experimental approaches, Carpenter and Stehr (2008) found that mandatory seat belt laws significantly reduce traffic fatalities and serious injuries due to fatal crashes by 8 and 9%, respectively. According to Elvik et al. (2009), the use of seat belts reduces the probability of being killed by 40–50% for drivers and passengers in front seats and by about 25% for those in rear seats. The effect of seat belt use on serious injuries is almost as great, whereas the effect on slight injuries is somewhat smaller, with a percentage of around 20–30%. The mandatory use of seat belts has, on average, led to a decrease of around 10–15% in the number of car occupants killed or seriously injured. The total number of car occupants injured is reduced by about 10%. The more seat belts are used, the greater the decrease in the number of injured persons in light vehicles. The greatest decrease observed is around 15–20%.

The effect of implementing the 1978 Seat Belt Rules (rules in which it is compulsory for drivers to fasten their seat belts before driving) on the rate of fatalities due to reckless driving is investigated in this study. The ARIMA(8,1,0) model predicts that the rate of

car driver fatalities will increase by about 69% without the implementation of the 1978 Seat Belt Rules. In contrast, the transfer function-noise model predicts a decline of 58% ($p = 0.025$) in the year in which these rules are implemented. The reduction is found to be statistically significant at 5% level. The rate of vulnerable road user fatalities is also investigated to determine whether there is a risk compensation effect. The results show that there is a 6% decline ($p = 0.001$) in the rate of vulnerable road user fatalities in the year when the rules are imposed compared to a 1% decline a year before the rules are imposed.

The results also show that the reduction in the rate of vulnerable road user fatalities is 78% in the long term. In contrast, the ARIMA(1,1,2) model predicts that the rate of vulnerable road users fatalities will increase about 9% in the absence of the 1978 Seat Belt Rules. Hence, based on estimates produced by the transfer function-noise model, it is expected that risk compensation effect will not occur due to the implementation of the 1989 Seat Belt Rules. This is in agreement with the findings of Evans and Graham (1991), and Cohen and Einav (2003), in which there is no positive effect of seat belt usage on the rate of non-occupant fatalities. It is expected that there will be fewer driver fatalities in Malaysia since the enforcement of the seat belt rules at the national level is rated 4 based on a scale of 0–10 (WHO, 2013).

5.2.1.3 Road Transport Act 1987 for Drunk Driving

The Road Traffic Ordinance 1958 consisting of enactments for traffic operations were revoked by Act 333 Road Transport Act 1987. This new Act came into force on 1st January 1988 and includes tougher penalties for driving while under the influence of intoxicating liquor or drugs. The Act states that any person who, when driving or attempting to drive a motor vehicle on a road or any other public place, is under the

influence of intoxicating liquor or a drug to such an extent that the person is incapable of having proper control of the vehicle, shall be guilty of an offence and shall on conviction be liable to a fine not exceeding two thousand ringgit (previously one thousand only) or to imprisonment for a term not exceeding six months and, in the case of a second or subsequent conviction, to a fine not exceeding four thousand ringgit (previously two thousand only) or to imprisonment for a term not exceeding one year or to both.

The Road Transport Act 1987 for drunk driving was superseded in 1994 by Act A878. This revision includes the limit of alcohol concentration in breath, blood and urine. In this study, the impact of the drunk driving legislation on road traffic injuries is investigated by evaluating the changes in the rate of accidents due to the influence of alcohol per 10,000 registered vehicles from 1969 to 1996. Year 1988 is used as reference since this is the year when the Road Transport Act 1987 was introduced extensively to the public.

In the absence of the 1987 Road Transport Act, the ARIMA(2,0,0) model estimates that the rate of accidents due to the influence of alcohol will increase by 7% in 1988. The execution of the road transport act with tougher penalties, however, has reduced the rate of accidents by about 63% in 1988 (the year in which a more stringent Act was implemented). The reduction is, however, not statistically significant since the ARIMA model predicts a declining trend during the period even without the 1987 Road Transport Act. In addition, it is possible to attain a lower rate of fatalities in Malaysia since the enforcement of drunk driving laws at the national level is rated 4 in 2010 based on a scale of 0–10 (WHO, 2013). Researchers at MIROS conducted a study to determine the incidence of driving under the influence of substance use (alcohol and drugs) among fatally injured drivers involved in road traffic crashes. The data were based on post-mortem files retrieved from the Department of Forensic Science, Kuala Lumpur Hospital. The findings revealed that 23.3% of fatal drivers were tested positive for alcohol, 11%

were tested positive for drugs and 2.3% were tested positive for both drugs and alcohol (Norlen et al., 2012).

The declining trend is consistent with the effect of implementing laws on drunk driving in many countries. A number of studies in the 1980s suggest that alcohol consumption will increase the frequency of fatal accidents (Hakim et al., 1991). Hilton (1984) evaluated the deterrent impact of a new set of strict laws governing drinking-driving offences in the State of California which came into effect in 1982. The study revealed that the number of fatal accidents reduces by 12.9% relative to the pre-intervention mean. Neustrom and Norton (1993) analysed the impact of stringent drunk driving laws in Louisiana using interrupted time series analysis, and discovered that these laws result in a 20% decrease in the mean of the pre-intervention night series.

In addition, Japan passed a new road traffic law in June 2002 that is intended to reduce alcohol-impaired driving by decreasing the permissible blood alcohol levels and imposing tougher penalties. Nagata et al. (2008) analysed the effect of a new road traffic law and found that the rate of alcohol-related traffic fatalities per billion kilometres driven decreases by 38% in the post-law period. A significant improvement in traffic safety becomes apparent with the implementation of a new law in the Republic of Serbia concerning driving under the influence of alcohol. It is found that there is a 26% decrease in the total number of traffic accidents per year after the first year of its implementation (2010) compared to 2009 (Zivkovic et al., 2013). In 2008, a law was enacted in Brazil that stipulates the concentration limit of zero blood alcohol for motorists. This law became known at the outlets as 'Prohibition'. Campos et al. (2013) analysed this law and the results indicate that there is a 45% decrease in driver behaviour after enactment of the new traffic law along with positive breathalyser results.

5.2.1.4 Speed Limit Rules in 1989

Earlier studies on speed limits have shown that speed limits have a marked effect in reducing the number of fatal accidents. However, the effect is less pronounced on slight or damage-only accidents (Smeed, 1961). An energy crisis in the US due to Arab oil embargo took place in 1973–1974. In response to the crisis, a large number of states have enacted a lower speed limit in order to decrease fuel consumption. Robertson (1977) found that some states experienced a reduction in fatalities up to 80% due to the combined effect of gasoline shortages and reduced maximum speed limits. According to Meier and Morgan (1980), the overwhelming percentage of decline in traffic fatalities from 1973 to 1974 and the ongoing decline is attributed to the 55 mph speed limit. Decreased travel is a minor factor compared to the estimated speed since fuel consumption is reduced by only 2.9%.

However, the viewpoint that lowering speed limits will reduce the number of injuries has been opposed by some researchers (for e.g. Lave, 1985; Lave & Elias, 1994; McCarthy, 2000; Balkin & Ord, 2001). When most cars travel at about the same speed (regardless high or low speeds), the rate of fatalities is low presumably because the probability of collisions will be low. These researchers perceive that the variability of vehicle speeds results in a higher number of fatalities rather than the speed itself. Thus, it is believed that reducing the variability of vehicle speeds should be the main focus because slow drivers are as much a public hazard as fast drivers (Lave, 1985). Balkin and Ord (2001) used stochastic structural equation modelling to predict the effect of speed limit changes on the number of fatal crashes on both urban and rural interstate highways in the US. Based on the results, the belief that higher vehicle speeds are indicative of a higher number of fatalities cannot be universally supported.

According to Baruya (1998), earlier studies have proposed that the cause of road accidents is the variance of speed, rather than the mean speed itself. Solomon's U-shaped hypothesis claims that both slow and fast vehicle speeds relative to the mean traffic speed is the cause for all types of accidents. However, recent studies do not provide evidence that slow vehicle speeds are in any way associated with an increase in personal injury accidents. According to Richardson and Shaw (2009), it is widely accepted that higher speeds will increase the severity of crashes, but the evidence on the impact of higher speed limits on US traffic fatality rates is somewhat mixed. Several studies have shown that the adoption of the 55 mph speed limit in 1974 reduces the number of traffic fatalities – however, the results are not as clear when the speed limits are increased in the 1990s. Feng (2001) concluded that speed and safety issues are interconnected to one another that it is difficult to distinguish whether a specific factor affects safety or speed. Most of the time, a factor influences both safety and speed. These factors include environmental conditions, driver behaviour and speed limits. Wang et al. (2013) suggested that more research is required to determine the exact relationship between speed and safety.

In response to the inconsistencies in the findings regarding the variability in speed limits, Kweon and Kockelman (2004) deduced that the inconsistencies are largely attributed to different methodologies and data settings, including temporal and spatial aggregation. In general, however, it has been shown that an increase in speed limits results in a higher number of fatalities per mile driven. Illegal speeding is found to be a major factor in casualty crashes and it can be expected that even a slight reduction in vehicle speeds will significantly reduce the frequency of casualty crashes in urban areas (Kloeden et al., 2002). According to Nilsson (2004), the reduction in speed limits from 110 to 90 km/h will decrease the number of injurious and fatal accidents by 25 and 52%, respectively. Friedman et al. (2007) discovered that a slight increase in the speed limit (6 mph) results in an immediate, substantial and persistent increase in the number of road

deaths – an increase of 12.7% on interurban roads. According to Elvik et al. (2009), a full compliance with the speed limit regulation reduces the number of fatal and injurious accidents by 15 and 9%, respectively.

More recent studies have shown that speed enforcement through the implementation of speed cameras has led to a substantial reduction in road crashes and fatalities. The introduction of speed cameras in Great Britain decreases the number of injurious accidents by 31.26%, which clearly indicates that the use of speed cameras helps in decreasing the number of accidents (Hess & Polak, 2003). The installation of fixed cameras and the deployment of mobile cameras results in a substantial reduction in speed, whereby the average speed decreases by 10 km/h in 2008, with 10% of the drivers driving at 10 km/h higher than the speed limit. A sharp decline in the number of deaths due to road traffic accidents was recorded between 2001 and 2008, with a value of 48% (Chapelon & Lassarre, 2010). In Guangdong Province in China, the number of motorway fatalities due to speeding decreases by 32.5% in 2005 relative to year 2004 upon the implementation of speed detection equipment (He et al., 2013).

In Malaysia, in exercising the power conferred in the Road Traffic Ordinance 1958, the Minister of Transport enacted the Motor Vehicles (Speed Limit) Rules 1959 which came into force on 1st July 1959. The maximum speed limit is specified as 40 mph (64 km/h) for passenger vehicles and 35 mph (56 km/h) for goods vehicles fitted with pneumatic tyres on all wheels of the vehicle. The Minister of Transport then enforced the Motor Vehicles (Speed Limit) Rules 1989 which revoked the speed limits enacted in 1959. According to this new set of rules, the maximum speed limit is 110 and 90 km/h for passenger and goods vehicles, respectively. It is apparent that the new speed limits almost doubled from the previous values. The significant changes in the maximum speed limit are examined in order to determine its impact on road traffic accidents. The analysis

conducted in this study reveals that the enactment of the Motor Vehicles (Speed Limit) Rules 1989 reduces the rate of motor vehicle accidents by approximately 6% in the year in which the law was enacted. However, the reduction in the rate of motor vehicle accidents is not statistically significant.

In addition, it is found that the law increases the rate of accidents by about 6% in the long term. The reduction during the first year of implementing the law may be attributed to the fact that the road users are unclear regarding the details of the new law. However, once the road users are aware regarding the increase in the maximum speed limits, the rate of road traffic fatalities follows the worldwide trend: increasing the maximum speed limits will increase the rate of fatalities.

In terms of advanced enforcement for speed limit offenders, the use of red light surveillance cameras and speed cameras came into effect in 1993 and 1994, respectively (PDRM, 1995). Speed enforcement, however, may not be as comprehensive as has been hoped until recent years. It is found that the enforcement of the national speed limit law is rated 5 in 2010 on a scale of 0–10 (WHO, 2013). Recently, the Government of Malaysia has implemented a new advanced system known as the Automated Enforcement System (AES). The AES is able to detect those who drive above the speed limit as well as those running a red light. This system came into effect on 22nd September 2012. These cameras will be installed at 831 black spots and implemented in stages.

5.2.1.5 Motorcycle Daytime Running Headlight Regulation in 1992

In the literature, the daytime headlight laws were first enacted in four states in the US, (Montana, Arkansas, Indiana, and Oregon) in 1967 (Muller, 1982). It is observed that the findings in previous studies are rather inconclusive since the early years of investigating the effects of running motorcycle headlights at daytime. One of the notable studies that

proved that running daylights is effective is the work of Janoff and Cassel (1971), in which they analysed motorcycle accident records of four states with daytime motorcycle headlight laws (Indiana, Montana, Oregon, and Wisconsin) and four control states. They found that the total number of motorcycle accidents in the four states with daytime headlight laws is reduced by 41.3 and 37.5% at daytime and night time, respectively. This gives a difference of 3.8%, which indicates that the effect of daytime headlight operation is significant at the 0.1 level.

According to Hurt et al. (1981), the use of headlights during the day is a powerful and effective way of reducing accidents by making the presence of motorcycles more conspicuous in traffic. Conversely, Lund (1979) used the odds-ratio test to analyse the effect of the enactment of motorcycle daytime headlight laws in Denmark in 1977, and found that there is an increase in the number of accidents, rather than a decrease. Muller (1982) examined the effectiveness of the laws which mandate the use of motorcycle headlights at daytime by comparing the proportion of fatal, front, and side-angle collisions at daytime between states with and without such laws. The comparison is based on all motorcycle fatalities recorded by NHTSA within 1975–1980. The analysis also shows that there is a statistically insignificant difference between states with and without such laws, suggesting that daytime headlight laws are ineffective.

Zador (1983), however, rejected the substantive conclusion of Muller's (1982) study on the basis of alleged methodological problems and flawed assumptions with the research. To support his argument, Zador (1985) analysed fatal motorcycle crashes in the US from 1975 to 1983. It is found that approximately 600 daytime crashes are prevented by the implementation of motorcycle headlight laws in 14 states during the study period, which corresponds to a reduction of 13% in the number of fatal daytime crashes. Muller (1985) in turn, refuted Zador's (1985) argument and claimed that the results reported by

Zador may be overestimated or spurious. Muller used time series analysis to support his findings.

Preliminary analysis on the short-term impact of running headlights intervention by Radin Umar et al. (1996) reveals that there has been a significant drop (29%) in conspicuous-related motorcycle accidents in Seremban and Shah Alam, Malaysia. Bijleveld (1997) analysed the effect of running headlights during daytime for motorcycles in the European Union, particularly Austria, in which a new law was introduced in 1982. Using a generalised linear model, the results showed that this law reduces the number of motorcyclist victims involved in multiple accidents at daytime by roughly 16%. In a review pertaining to the use of motorcycle and motorcar daytime headlights, Perlot and Prower (2003) summarised that there is no satisfactory scientific evidence on the reduction of accidents after 35 years of implementing the first motorcycle daytime lights laws in the US.

Elvik et al. (2009) reviewed studies on running headlights at daytime and discovered something contradictory. There are two studies related to the effects of running headlights at daytime for mopeds and motorcycles. The aim of the first study is to determine the effect of running headlights at daytime on the accident rate of each motorcycle (individual effect). The aim of the second study, however, is to investigate the effect of running headlights at daytime on the total number of accidents in a country in which running headlights during the day is mandatory for motorcycles. It is found that the mandatory use of headlights on mopeds and motorcycles at daytime reduces the number of multi-vehicle accidents by roughly 7% ($\pm 2\%$).

The findings of this study, however, show that there is no reduction in the national rate of motorcycle accidents following the Motorcycle Daytime Running Headlight Regulation in 1992. Based on the intervention analysis model, the rate of motorcycle

accidents increases about 6% a year after the regulation was implemented. However, the increase is not statistically significant. The rate of accidents continues to increase several years after the regulation was imposed. In fact, the rate of motorcycle accidents shows a declining trend before the regulation came into effect in 1992. Again, the percentage of motorcyclists who switched on headlights at daytime and the level of enforcement conducted following the regulation may be the reason why there is no positive effect. For a smaller jurisdiction (Seremban and Shah Alam), Radin Umar et al. (1996) observed that there is a significant decline due to the implementation of the daytime headlight regulation. It shall be noted that these areas are urban areas which are equipped with better road safety campaigns and enforcement. Hence, it can be concluded that the Motorcycle Daytime Running Headlight Regulation in 1992 is not an effective road safety measure to reduce the rate of fatalities in Malaysia at that time.

5.2.1.6 Specifications for Protective Helmets (MS 1:1996)

Motorcycle helmets were initially no more than leather bonnets – first used in racing and usually worn with goggles. These skull caps were adapted from earlier aviators whose primary purpose was to attain comfort and therefore, there was almost no protection provided to the head. Hence, crash helmets did not exist at the time. The padding of modern helmets, which is basically a resilient closed cell rubber foam placed within the interior of the shell in order to dissipate impact energy effectively was introduced by Turner and Havey (1953). However, this design was rather heavy. The helmet was then improved by Roth and Lombard (1953) who developed modern helmets as they are known today. Recently, modern helmets are capable of distributing impact loads over a large area of the head and reducing the total force on the motorcyclist's head as much as possible (Fernandes & Sousa, 2013).

Motorcycle's riders are over 30 times more likely to die in a traffic crash than car occupants (Lin & Kraus, 2008). However, according to Elvik et al. (2009), the injury rate for moped riders and motorcyclists in traffic is 12–15 times as high as that for car drivers. Motorcyclists are less protected against road accidents than the users of other vehicles because crash helmets are their only most effective means of protection, whereas car occupants are protected by safety belts, airbags as well as the body structure of the car (Fernandes & Sousa, 2013). Moreover, head injury is one of the most frequent injuries resulting from motorcycle accidents. NHTSA (1998a) estimated that motorcycle helmets are 36% effective in preventing deaths and according to Deutermann (2005), this effectiveness has increased over the years which is possibly due to improvements in helmet designs and materials. Motorcycle helmets have long been recognized as an effective countermeasure to prevent head and brain injuries among collision-involved motorcycle riders (Tsui et al., 2013). Efforts made to improve helmet designs for motorcyclists are likely to improve their overall safety.

Tsai et al. (1995) investigated the effectiveness of various types of helmets on the prevention of head injuries of motorcyclists. They used logistic regression analysis to determine the roles of the following variables in predicting the risk of head injuries: age, sex, riding position, weather, place of accident, type of helmet, type of motorcycle, and the status of helmet wearing. The results showed that the risk of head injuries among motorcyclists is reduced significantly by wearing a full-face helmet rather than a full-coverage or partial-coverage helmet. Nevertheless, the issue of non-standard helmets is still prevalent even in the US, as reported by Tsui et al. (2013). They conducted a survey in San Francisco Bay–area in 2008 and discovered that 15% of the motorcyclists sometimes or often wear non-standard helmets. Ackaah et al. (2013) conducted a cross-sectional survey of helmet-wearing motorcyclists, a market survey of helmet prices, and a review of current legislation and enforcement of policies and practices. Nine countries

participated in this study: China, Ghana, India, Malaysia, Mexico, Nigeria, Pakistan, Thailand and Vietnam. The data were collected between May 2008 and May 2009. A total of 5563 helmet-wearing motorcyclists were observed and the interviewers judged that 49% of the helmets involved in the survey were likely to be non-standard helmets.

In Malaysia, the first standard for motorcycle safety helmet (MS 1:1969) was released in 1969 by the Standards Institution of Malaysia (SIM). This standard was then superseded by MS 1:1996. There are major changes made in the second revision of the helmet standard. The methods for testing impact resistance, penetration and strength of retention system have been revised in line with current international practice. In addition, in the revised standards, the tri-axial acceleration is taken as the headform acceleration in the impact energy test. The third revision of the safety helmet standards was released by the Standards and Industrial Research Institute of Malaysia (SIRIM) in 2011. According to MIROS (2014), this third revision adopts a significant amount of content from UN R22 – however, the penetration test is retained since this test was not included in the UN R22. This version also includes a visor standard (which was previously a stand-alone standard) in Part 2 of MS1.

The transfer function-noise model developed in this study predicts that the implementation of the second revision of the helmet standard will significantly reduce the rate of motorcycle fatalities related to head injuries. There is about 31% reduction ($p < 0.0001$) in the rate of fatalities related to head injuries in 1997, a year after the standard came into effect. This is consistent with the findings of other researchers, whereby improved helmet standards will reduce the risk of head injuries. It is believed that this reduction can be increased since there is a relatively high number of motorcyclists who are still using non-standard helmets. Ackaah et al. (2013) found from interviews that 35% of the helmets worn in Malaysia are non-standard helmets. Rabihah (2013) studied the

compliance of motorcyclists towards helmet use during the enforcement period in two areas in Selangor, Malaysia. Even though there is high compliance towards helmet use, the rate of proper use is rather dismal, with a range of 43.8–46.3%. A low rate of proper helmet use indicates the need for improvements in enforcement activities and more importantly, motorcyclists do not understand the philosophy underlying the laws of helmet use. It has been highlighted that there is need to enhance enforcement approaches in Malaysia as well as proper advocacy programmes to mitigate this situation.

5.2.1.7 Road Safety Programmes in 1997

Integrated road safety programmes have been drafted and implemented as a follow-up to the National Road Safety Target in 1996. According to Law et al. (2005), the integrated road safety programmes in 1997 include targeted televised motorcycle safety campaigns, more stringent traffic legislation, national accident black spot programmes, road safety auditing, construction of new motorcycle lanes and better protection for motorcyclists. In this study, the rate of road traffic fatalities at pre-implementation and post-implementation stages of the programmes are examined in order to determine the effectiveness of such programmes. If the integrated road safety programmes are excluded, the ARIMA(6,1,6) model predicts that there is a negligible decrease in the rate of fatalities, with a value of 0.5%. If the integrated road safety programmes are included, the transfer function-noise model predicts that there will be a decrease in the rate of fatalities after a year of implementing the programmes, with a value of 6%. This decrease, however, is not statistically significant. The model predicts that the decrease in the rate of fatalities is 9% in the long term. This achievement is considered ‘moderate’ compared with the road safety achievements in high-income nations.

The national safety targets or plans that have been implemented in other countries are described as follows. In the National Road Safety Strategy 2001–2010, the Australian Transport Council sets a target of reducing the number of fatalities per 100,000 population by 40% by year 2010. This target is to be achieved in four major categories of road safety measures. Each of these categories is expected to contribute to the overall 40% reduction in the number of fatalities. The four categories of road safety measures are ‘improving road user behaviour’ (9%), ‘improving vehicle occupant protection’ (10%), ‘using new technology to reduce human errors’ (2%) and ‘improving the safety of roads’ (19%) (Loo et al., 2005). Based on this 10-year target, the targeted annual reduction is 4% on average.

According to Koornstra (2007), road reconstruction, intensified police enforcement and vehicle safety measures result in a decrease in fatalities of approximately 85, 70, and 40% for each road safety measure in high-income countries. If these measures are combined altogether, it can potentially save a large percentage of road traffic fatalities because saving $100 \times [1 - (1 - 0.85) \times (1 - 0.7) \times (1 - 0.4)]$ equals to 96% fatalities (after correcting for fatalities already saved by other safety measures). However, road reconstruction requires a considerable amount of time and must be financed by long-term investment plans. It shall be noted, however, that the annual decrease in fatalities that can be achieved is not detailed in this estimation. Thus, the level of implementation plays a key role in this case.

Lu (2007) reported that the annual number of fatalities in Netherlands decreases substantially by 7.2% during the implementation period of Phase I of the Dutch road infrastructure redesign programme (1998–2002). The annual number of fatalities decreases dramatically within 2002–2006 with a percentage of 23.9%. According to Elvik et al. (2009), based on the study of 22 quantified targets set by national governments and 13 quantified targets set by regional or local governments on road safety performance, it

can be concluded that it is not possible to evaluate the effects of quantified targets on road safety performance in a sufficiently rigorous manner in order to draw a conclusion regarding their effects. If the tendencies present in the data are taken as indicative of real effects, then a quantified target is associated with only a slight improvement in road safety performance, amounting to a net reduction in the number of road accident fatalities of about 0.8% per year. The largest improvement in road safety performance is associated with ambitious, long-term targets set by national governments. In addition, the possible effects of 100% compliance with typical road traffic legislations on the number of people killed or injured is a reduction of about 48% and 27%, respectively.

5.2.1.8 Integrated Road Safety Operations (*Ops Sikap*) in 2001

Since the early years of road safety studies, researchers have emphasized on the three E's in order to improve road safety: (1) education, (2) enforcement and (3) engineering. According to Votey (1984), the increase in the alcohol consumption is associated with a higher incidence of accidents which in turn, increases the level of law enforcement. This leads to a greater number of sanctions which decreases the number of fatal accidents as well as accidents which will result in serious injuries. Haque and Cameron (1989) analysed the effect of zero BAC legislation introduced in Victoria, and found that there is no significant reduction in accidents which is probably due to the absence of specific enforcement procedures and mass media publicity campaigns.

Mohan (2004) summarised the following key points in his review regarding the effect of enforcement on road safety: (a) Most attempts at enforcing road traffic legislation will not have any lasting effects, either on road user behaviour or on crashes unless the enforcement is continuous and widespread. This level of enforcement is expensive and difficult to sustain in most situations. (b) Imposing stricter penalties (in the form of higher

finer or longer prison sentences) does not affect road-user behaviour significantly and imposing stricter penalties also reduces the level of enforcement. (c) Increased normal, stationary speed enforcement is cost-effective in most cases. Automatic speed enforcement with cameras seems to be more efficient. (d) There is no evidence that proves that mobile traffic enforcement with patrol cars is cost-effective.

According to Kim et al. (2006), it is believed that the enforcement of a road safety policy in Korea at the onset of the policy results in an abrupt and instant (i.e. no time delay) effect. It has been shown that the reduction in fatalities is more than 70% in the US after one decade of intensive enforcement, as well as the use of appropriate safety devices and limit laws (Koornstra, 2007). Soole (2009) investigated the impact of police speed enforcement methods on self-reported speeding behaviour. The results indicate that marked patrol vehicles parked on the side of the road are most effective to curb speeding behaviour on freeways (62.7%) and school zones (56.3%), but only somewhat effective on urban roads (49.4%). According to Antić et al. (2011), augmenting enforcement means more severe penalties and frequent control, which in turn will decrease the number of traffic accidents and violations of road safety. This will have a positive effect on increasing road traffic safety.

In anticipation of the hike in accident tolls during festive seasons, a number of government agencies have been established under the Ministry of Works, Ministry of Transport, Royal Malaysian Police and the Road Transportation Department since 2001. These government agencies work collaboratively on integrated road safety operations in order to improve road traffic safety in a joint programme known as '*Ops Sikap*'. The operations comprise various approaches including engineering, media campaigns and enforcement. *Ops Sikap* is carried out during festive seasons such as *Hari Raya Aidilfitri* and Chinese New Year which is celebrated by Muslims and Chinese, respectively. The

duration of the *Ops Sikap* is 15 days. The daily traffic volume increases during these festive seasons – it has been shown that the daily traffic volumes on several expressways increased from 6 to 75% in 2011.

Jamilah (2012) evaluated the effectiveness of *Ops Sikap* using data collected from 2009 to 2011. It is found that *Ops Sikap* neither decrease nor increase the total number of fatalities and number of motorcycle fatalities during *Ops Raya* and *Ops CNY* in 2009, 2010 and 2011 except for the total number fatalities during *Ops CNY* in 2010 and 2011. Yacoob et al. (2011) investigated the effect of implementing *Ops Sikap* and the results showed that there is a decrease in the number of road accidents in Malaysia. However, a significant reduction is only observed after the implementation of *Ops Sikap VII* and *Ops Sikap IIV* conducted during *Hari Raya Aidilfitri* in 2004 and 2007, respectively.

One of the goals of this study is to examine the effect (in percentage) of the integrated road safety operations from 2001 to 2005. Since the road safety plan was launched in 2006, the observations are limited up to year 2005 due to the fact that the safety measures implemented from year 2006 onwards may influence the rate of fatalities. According to the ARIMA(12,1,4) model, the rate of road traffic fatalities will increase by roughly 5% in 2001 in the absence of *Ops Sikap*. However, based on statistics of the actual *Ops Sikap* conducted in 2001, the rate of fatalities is reduced by about 10%. The reduction in the rate of fatalities is found to be statistically significant. Consistent operations in the long term reduce the rate of fatalities up to 18%. The results indicate that *Ops Sikap* is an effective safety measure to reduce fatalities in the early years of implementation.

5.2.1.9 Road Safety Plan 2006–2010

The effectiveness of road safety plans in Scandinavian countries has been highlighted by Elvik (2010a). The outcomes in Denmark, Finland and Sweden are close to the targets set in the road safety programmes during the first few years of implementation, but the progress slows down thereafter. The actual number of road accident fatalities exceeds the number of fatalities targeted in the road safety programmes at the end of the period. In contrast, the number of fatalities develops more favourably in Norway than the targeted value stated in the National Transport Plan within the following period (2002–2011). Rizzi (2011) conducted a study to evaluate the effectiveness of the road safety plan in Chile by estimating the benefits of the road safety measures. It is concluded that implementing the eight road safety measures specified in the plan reduces the number of fatalities by 42% (95% CI: 22% - 57%). The impact of these eight measures seems impressive, but most of the benefits are obtained from speed control and reducing the legal speed limits at night.

The Road Safety Plan 2006–2010 was launched in Malaysia in order to be at par with developed countries. The plan consists of nine strategies and 52 programmes. The first strategy is focused on road safety education with the aim of raising public awareness towards road safety. The second strategy is focused on enforcement in order to reduce the number of accidents resulting from reckless driving through technological expertise. The third strategy makes use of an engineering approach, comprising programmes such as rehabilitation of accident prone areas and development of exclusive lanes for motorcycles. The fourth strategy is focused on increasing the public's involvement in these activities whereas the fifth strategy is focused on encouraging the use of public transport. The sixth strategy is focused on other critical sectors whereas the seventh strategy is focused on high-risk users. The eighth strategy is focused on reviewing and

improving road safety rules whereas the ninth strategy is focused on promoting the sharing of funds in order to finance road safety programmes.

The following road safety targets are set in the plan: (1) 2.0 deaths per 10,000 registered vehicles, (2) 10 deaths per 100,000 population and (3) 10 deaths per billion of vehicle kilometres travelled by 2010. However, these targets are underachieved since the rate of fatalities in Malaysia is 3.4 per 10,000 registered vehicles and 24.3 per 100,000 population in year 2010. These rates of fatalities are still considerably higher than the targets set in the plan. In order to achieve a target of 2.0 deaths per 10,000 registered vehicles, there is a need to achieve a reduction of 52.4% deaths from 2005 to 2010 (Rohayu et al., 2012). However, based on the prediction of the transfer function-noise model developed in this study, the road safety programmes stipulated in the national plan can only reduce the rate of fatalities by approximately 7%. However, the rate of fatalities in the long term is only 9%, which is far from the targeted reduction of 52.4%. More importantly, the reduction in the rate of the fatalities due to implementation of the plan is found to be statistically significant.

5.2.2 Which is the best time series model among the three models developed in this study in order to predict the rate of fatalities in Malaysia?

There are two types of time series models developed in this study to forecast the rate of road traffic fatalities in Malaysia: univariate and multivariate models. It shall be highlighted that two univariate models (ARIMA(2,1,0) and transfer function-noise models) and one multivariate model (state-space model) are developed for forecasting purposes. The MAPE values are used to determine the extent to which the forecast values have deviated from the actual values, and they can be used to assess the accuracy of the models which have been developed using different techniques and data series. Hence, the

best model to forecast the rate of road traffic fatalities in Malaysia is determined based on the MAPE values. In this study, the MAPE value of the training data (from year 1985 to 2012) for transfer function-noise model, ARIMA(2,1,0) model and state-space model is 1.1%, 3.0% and 3.4%, respectively. Whilst, for test data (year 2013 and 2014) is 0.6%, 6.1% and 4.5%, respectively. Hence, the transfer function-noise model is selected as the best model since it has the smallest error in forecasting the fatalities.

It is also found that the rate of fatalities forecasted by all of the three models shows a declining trend. However, the decrease in the rate of fatalities varies from one model to another, whereby the most significant decrease is obtained by the transfer function-noise model, followed by the state-space model. The ARIMA(2,1,0) model gives the slowest decrease in the rate of fatalities among all models. This leads to the following question: Which is the best model to forecast the rate of fatalities in the future? Based on the statistical results presented in Chapter 4, it can be inferred that all of the three models are suitable to forecast the rate of fatalities. However, it is worth mentioning that one should consider the level of enforcement of road safety legislations and other road safety measures since they play a vital role in influencing the number of fatalities in the future.

Basically, if the level of enforcement is as intensive as that conducted from 1997 to 2006, the rate of fatalities will follow the trend predicted by the transfer function-noise model. According to this model, the rate of fatalities will decrease significantly, whereby the absolute number of fatalities will reach its peak in 2011 and will begin to decrease from year 2012. In reality, however, the rate of fatalities still increases after 2012, which indicates that there may be a lack of enforcement and/or road safety measures in the last few years. In contrast, if the level of enforcement is not as intensive as in previous years, the rate of fatalities will follow the trend predicted by either the ARIMA(2,1,0) model or state-space model.

5.2.3 Is it acceptable to only use the descriptive/univariate model (without explanatory variables) which is practised extensively in Netherlands?

The number of fatalities in developed countries had reached its peak in 1970–1972 and has decreased since then. It is possible that this trend will also occur in developing countries including Malaysia. A robust and comprehensive model such as the DRAG model is required in order to predict the time at which the peak number of fatalities will occur. However, this model requires a large number of explanatory variables which may be impractical for developing countries due to limited availability of data. Based on the results obtained in this study, the univariate model is practical to forecast the number of fatalities in Malaysia. Nevertheless, the univariate model is suitable to forecast the rate of fatalities rather than the absolute number of fatalities. The models developed in this study confirm that the rate of fatalities follows a declining trend. In addition, the historical number of fatalities is always increasing, whereas the number of fatalities may decrease in the future. In short, the univariate model is practical to forecast the rate of fatalities rather than the absolute number of fatalities.

5.2.4 What are the significant explanatory variables which affect the number of fatalities?

In this study, preliminary screening is carried out on the variables that are expected to have a strong correlation with the number of fatalities, as suggested by other researchers. The trend of each variable is compared with the trend in the number of fatalities and available data in order to select the appropriate variables. Hence, 15 variables are selected in order to develop the forecasting models:

- 1) GDP per capita (constant 2005 US\$)
- 2) Hospital beds (per 1,000 people)
- 3) Population (total, in 100,000)
- 4) Male population (% of total)
- 5) Female population (% of total)
- 6) Population aged 15–24 years old (% of total)
- 7) Registered vehicles (in 10,000)
- 8) Registered motorcycles (in 10,000)
- 9) Registered buses (in 10,000)
- 10) Registered motorcycles (% of total)
- 11) Road sector energy consumption (kt of oil equivalent)
- 12) Road length (in 1,000 km)
- 13) Rural population (% of total)
- 14) Urban population (% of total)
- 15) Unemployment rate (in %)

Multiple regression analysis is used to identify the variables with a p -value less than 0.1. However, it shall be noted that these two variables (female population and urban population) are removed by the statistical software since there is a high correlation between these variables and other variables. Following this, multiple regression analysis is carried out for the second time with the remaining 13 variables in order to exclude variables with multi-collinearity. The significant explanatory variables obtained from the multiple regression analysis are as follows:

- (a) Hospital beds per 1,000 people (p -value = 0.000)
- (b) Percentage of registered motorcycles (p -value = 0.006)
- (c) Road length in 1,000 km (p -value = 0.040).

5.2.5 Is it possible for the number of fatalities to increase significantly on an ongoing basis considering the fact that most of road safety measures have been implemented?

In response to the question above, it is indeed possible for the number of fatalities to increase significantly even though most of the road safety measures have been implemented. This generally depends on the level of compliance with the road safety measures. The level of enforcement and widespread campaigns on road safety are also the key parameters for road safety achievement. Even though it is expected that the enactment of a number of legislations will partially decrease the rate of fatalities due to road traffic accidents, the trends show otherwise, indicating that the legislations do not have any effect. Moreover, the transfer function-noise model predicts a lower number of fatalities compared to the actual value for year 2012. In other words, the actual number of fatalities is higher than the predicted value based on historical data. This may be attributed to a lack of enforcement and promotion of road safety campaigns in the last few years.

5.2.6 Is the major decline of fatalities in 1997–1998 due to the initiatives and measures implemented in Malaysia? Or is it possible that this decline is a consequence of an unexpected regional economic crisis at the time which in turn reduces traffic exposure (vehicle-kilometres travelled or fuel consumption)?

Even though it is expected that 15 explanatory variables contribute to the number of fatalities, only three variables can be linked with the economic situation and traffic exposure, namely the number of registered vehicles, road sector energy consumption and road length. Based on the results of the multiple regression analysis, there is a significant

correlation between the percentage of registered motorcycles and the number of fatalities, as well as between the road length and the number of fatalities. It is observed that there are drastic changes in these variables which may significantly affect the reduction in the number of fatalities in 1997. However, the regional economic crisis in 1997 leads to a substantial reduction in the road sector energy consumption in 1998. A significant reduction in the number of fatalities is also recorded in 1998. Therefore, it can be summarized that the road safety programmes (including intensive enforcement and extensive campaigns in 1997) are the major factors which cause a decline in the number of fatalities. This is followed by a substantial decrease which is in line with the decrease in road sector energy consumption due to the economic downturn in 1998. The combination of these events leads to a decrease in the number of fatalities in 1997–1998.

5.2.7 Is it practical to include only the population and number of vehicles as the explanatory variables as had been done by MIROS?

The results of this study indicate that the forecasting model which includes only the population and number of vehicles as the explanatory variables is appropriate to predict the number of fatalities if there is less enforcement in the future. However, the probability of less intensive enforcement is low since it is expected that the Government of Malaysia will continue to implement a stringent enforcement system in the future. In addition, the forecasted number of vehicles (based on the fact that these vehicles are the major modes of transportation) should be interpreted with caution since the public transportation system such as MRT, LRT and buses is upgraded aggressively at present. This leads to a major shift from the use of private transport to public transport. In general, the various means of public transport are safer compared to privately owned vehicles. The number of fatalities predicted by all of the models in this study is substantially lower in 2015 and 2020 compared with the predicted values obtained by Rohayu et al. (2012). In fact, the

number of fatalities in year 2013 and 2014 is 6,915 and 6,674. Table 5.1 shows the forecasted number of fatalities produced by the models.

Table 5.1: Forecasted number of fatalities in 2015 and 2020 produced by the three different models developed in this study as well as MIROS

| Year | Actual & Forecasted number of road traffic fatalities | | | | |
|-------------|---|---------------------------------|--------|--------------------|-----------------------|
| | State-space model | Transfer function-noise (5,1,7) | Actual | ARIMA(2,1,0) model | ARIMA(0,1,1) by MIROS |
| 2013 | 7,032 | 6,859 | 6,915 | 7,122 | - |
| 2014 | 7,166 | 6,696 | 6,674 | 7,281 | - |
| 2015 | 7,301 | 6,488 | 6,706 | 7,487 | 8,760 |
| 2016 | 7,436 | 6,506 | 7,152 | 7,673 | - |
| 2020 | 7,972 | 6,610 | - | 8,442 | 10,716 |
| MAPE | 5.5% | 3.4% | | 7.8% | 30.6% |

5.2.8 Can the number of fatalities targeted in the Road Safety Plan 2014–2020 (death toll: 5,358) be achieved in year 2020?

Based on Table 5.1 above, the best model developed in this study (i.e. the transfer function-noise (5,1,7) model) forecasts that the number of fatalities is 6,610 in year 2020. Even though the number of fatalities can be reduced to 3,832 in 2020, the probability that the target will not be achieved is higher. Thus, in order to achieve the target in the Road Safety Plan 2014–2020, drastic road safety measures should be carried out compared to the previous road safety plan.

5.2.9 Has the risk compensation theory occurred?

There is no evidence that legislation of making seat belt use compulsory results in reckless behaviour and higher number of fatalities of vulnerable road users. Nevertheless, risk compensation may occur in Malaysia specifically for motorcyclists and motorcar occupants. The number of motorcyclists and motorcar occupants continues to increase over the years. There is no indication that the number of motorcycle and motorcar fatalities will decrease, even with the implementation of intensive enforcement in recent years. The production of vehicles with advanced safety features and enhanced road geometries has increased the driving intensity (reckless behaviour) of road users such as speeding.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

The characteristics of the road safety scenario in Malaysia have been determined and three time series models have been developed to predict the rate of road traffic fatalities in order to fulfil the research objectives.

6.1 Conclusions

The following conclusions are drawn based on the findings:

Characteristics of road safety in Malaysia

- The characteristics of the current road safety in Malaysia are: (1) the major victims of road traffic accidents are motorcyclists, (2) young adult drivers/riders aged 16–25 years make up the highest percentage of the total fatalities, with a value of 35%, and (3) the highest rate of fatal accidents per kilometre occurs at expressways.

The effectiveness of road safety measures in Malaysia

- The Safety Helmet Rules imposed in 1973 does not reduce the rate of motorcyclist fatalities, which may be attributed to the low percentage of helmet usage.
- The enactment of the Seat Belt Rules in 1978 results in a 58% decrease in the rate of motorcar driver fatalities in 1979. This decrease is found to be statistically significant.
- The implementation of the 1987 Road Transport Act with tougher penalties for drunk driving decreases the rate of accidents due to alcohol consumption by approximately 63%. However, this decrease is not statistically significant since

the rate of accidents due to alcohol consumption tends to decrease even without implementation of the Act.

- The enactment of the Speed Limit Rules in 1989 increases the rate of motor vehicle accidents by approximately 6%. However, this increase is not statistically significant.
- The Motorcycle Daytime Running Headlight Regulation in 1992 is not an effective road safety measure to reduce the rate of motorcycle accidents in Malaysia during the period of observation.
- The implementation of the Specifications for Protective Helmets (MS 1:1996), which is the second revision of the helmet standard, decreases the rate of motorcycle fatalities related to head injuries by 31% in 1997 – a year after the standard came into effect.
- Strategies to improve road safety in Malaysia should be focused on developing road safety measures for motorcyclists and motorcar occupants. It is believed that the appropriate safety road measures will greatly reduce the number and rate of fatalities for these road users, without compromising the efforts made to improve the safety of other road users.

The effectiveness of road safety programmes

- The implementation of the Road Safety Programmes in 1997 results in a 9% reduction in the rate of total road traffic fatalities over the long term.
- The implementation of the Integrated Road Safety Operations (*Ops Sikap*) since 2001 is found to be an effective road safety measure to reduce the rate of the total road traffic fatalities, particularly during the early years of its implementation.
- The implementation of the Road Safety Programmes in 1997 (which includes intensive level of enforcement and extensive road safety campaigns to promote

awareness among various road users) is found to be the major factor that causes the decrease in the rate of the total road traffic fatalities. This decrease occurs simultaneously with the substantial decrease in road sector energy consumption owing to the regional economic crisis in 1997. The combination of both events leads to a decrease in the number of fatalities within 1997–1998.

Road safety plan

- The implementation of the National Road Safety Plan 2006–2010 decreases the rate of the total road traffic fatalities by only 7% and 9% over the long term. This achievement is rather low compared to the reduction in the rate of total road traffic fatalities targeted in the plan, which is 52.4%. However, this decrease is found to be statistically significant.
- It is believed that the targeted rate of fatalities of 2.0 per 10,000 registered vehicles will be achieved by year 2023 by the implementation of the road safety programmes outlined in the National Road Safety Plan 2006–2010 coupled with intensive enforcement.
- In order to achieve the target in the Road Safety Plan of Malaysia 2014–2020, drastic road safety measures should be carried out compared to the previous road safety plan.

Result of the modelling

- If the level of enforcement is as intensive as that from 1997 to 2006, the trend of road traffic fatalities follows the trend predicted by the transfer function-noise model, which shows a rather sharp decline. However, if the level of enforcement is not as intensive, the trend of road traffic fatalities will follow the trend predicted by the ARIMA and state-space models.

- The univariate model is practical to forecast the rate of fatalities and not the absolute number of fatalities.
- The significant explanatory variables to forecast the number of fatalities are: (1) the number of hospital beds per 1,000 people, (2) the percentage of registered motorcycles and (3) road length.
- The number of road traffic fatalities in year 2020 forecasted by the state-space model is 7,972 (95% CI: 7,527–8,417), 6,610 (95% CI: 3,832–11,402) by the transfer function-noise (5,1,7) model and 8,442 (95% CI: 5,186–13,743) by the ARIMA(2,1,0) model, respectively. It shall be noted that the number of road traffic fatalities is predicted based on the assumption that the current efforts and programmes on road safety are sustained until year 2020.
- In general, the number of fatalities per 100,000 population is relatively high in Malaysia compared to other countries. The high number of registered vehicles results in a low rate of fatalities per registered vehicle. However, the rate of fatalities per registered vehicle is suitable to compare the road safety achievement in Malaysia with other countries since this parameter accounts for the effect of motorization level.
- The forecasting model which includes only the population and number of vehicles as the explanatory variables is appropriate to predict rate of the fatalities, provided that the enforcement is less intensive in the future. However, this situation is highly unlikely since the Government of Malaysia will implement a more stringent enforcement system in the future. It shall be highlighted that the forecasted number of vehicles (based on the assumption that these vehicles are the major modes of transportation) should be interpreted with caution, considering that aggressive efforts are currently made by the government to upgrade the public transportation system (particularly in urban areas) which includes MRT, LRT and

buses. This will lead to a major shift from the use of privately owned vehicles to public transport, which in turn will reduce the number and rate of road traffic fatalities since it has been proven that the use of public transport is safer than driving one's own vehicle.

Risk compensation in Malaysia

- Risk compensation may occur in Malaysia specifically for motorcycle and car occupants. The production of motor vehicles advanced safety features and ongoing improvements in road geometries has increases the driving intensity (reckless behaviour) of road users such as speeding.

6.2 Recommendations

Even though the objectives of this study have been met, there are still a number of areas which can be worked on in order to improve the accuracy of the forecasts and more importantly, to improve the road safety scenario in Malaysia. Some recommendations for future work are highlighted as follows:

- It is more desirable if the analyst acquires data on road traffic fatalities and injuries that are published on a monthly basis (if possible) in order to improve the accuracy of forecasts, particularly if time series analysis is used.
- It is essential for the analyst to acquire data on the road vehicle-kilometres travelled when making comparisons in road safety.
- Safety education is crucial since young adult road users aged 16–25 years account for 35% of the total number of driver/rider fatalities.
- Emphasis should be given on improving the safety provisions for motorcyclists.

- Emphasis should be given on post-crash countermeasures since the number of hospital beds is one of the significant factors associated with road traffic fatalities.
- Advanced public transport facilities are required to attract more motorcyclists and motorcar occupants to use such facilities. This will help reduce the number of vehicles travelling on the roads, which in turn, reduces the risk of road traffic fatalities.
- A more stringent enforcement system is needed in order to achieve the targets outlined in the road safety plan.

6.3 Contribution of the study

This study test the rate of fatalities targeted in the 2014 Road Safety Plan and found may not be achievable by year 2020. The results of the analysis may assist the parties involved in road safety (e.g. road transport authorities, public health agencies, vehicle producers, education sector, police and NGOs) to anticipate the trend of road traffic fatalities in the future and therefore, the necessary measures can be taken if the predicted death toll is beyond expectation. The results of this study can be used as a basis for comparison with the predicted data released by MIROS.

The significant explanatory variables that affect the number of fatalities are obtained in this study, which can be a comparison for the similar research. This study also reveals that the rate of fatalities will follow the trend predicted by the transfer function-noise model that is selected as the best model to forecast the number of fatalities in Malaysia. Moreover, it is shown that the forecasting of the number of fatalities using the same data with different software could result in significant difference of the fatalities.

6.4 Limitation of the study

In this study, it is deemed necessary to use the vehicle-kilometre travel (VKT) as the unit of exposure to risk. However, there are no comprehensive surveys on VKT in the early years, the surveys on VKT for all states in Malaysia were conducted by MIROS since year 2007. Therefore, the fatalities per 10,000 vehicle is used in this study.

6.5 Recommendation for future research

It is recommended to relate and forecast the number of road traffic fatalities based on the black spot treatment programme through road safety auditing that has been proven to be successful in reducing the number of fatal accidents. In this study, difficulties arise, when an attempt is made to detail how many black spots remained as since the figures change constantly due to the occurrence of new accidents. For this reason, predicting the total number of fatalities in Malaysia based on the number of remaining black spots as the independent variable is rather impractical at the moment.

REFERENCES

- Abbas, A. K., Hefny, A. F., & Abu-Zidan, F. M. (2012). Does wearing helmets reduce motorcycle-related death? A global evaluation. *Accident Analysis and Prevention*, 49, 249–52. doi:10.1016/j.aap.2011.09.033.
- Abdel-Aty, M. A., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention*, 32(5), 633–42.
- Abraham, B., & Ledolter, J. (2005). *Statistical methods for forecasting*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Ackaah, W., Afukaar, F., Agyemang, W., Thuy Anh, T., Hejar, A. R., Abdul, G., ... Yu, J. (2013). The use of non-standard motorcycle helmets in low- and middle-income countries: a multicentre study. *Injury Prevention: Journal of the International Society for Child and Adolescent Injury Prevention*, 19(3), 158–63. doi:10.1136/injuryprev-2012-040348.
- Adams, J. G. U. (1985). *Risk and freedom: the record of road safety regulation*. Cardiff: Transport Publishing Projects.
- Adams, J. G. U. (1987). Smeed's Law: Some further thoughts, *Traffic Engineering and Control*, 28(2), 70–3.
- Adams, J. G. U. (1994). Seat belt legislation: the evidence revisited. *Safety Science*, 18(2), 135–2. doi:10.1016/0925-7535(94)90022-1.
- Adams, J.G.U. (1982). The efficacy of seat belt legislation. *SAE Transactions*, 2824–38.
- Akaike, H. (1976). This is a chapter. In Mehra, R., and Lainiotis, D. G. (Ed.), *Advances and case studies in system identification: canonical correlations analysis of time series and the use of an information criterion*. New York: Academic Press.
- Akloweg, Y., Hayshi, Y., & Kato, H. (2011). The effect of used cars on African road traffic accidents: a case study of Addis Ababa, Ethiopia. *International Journal of Urban Sciences*, 15(1), 61–69. doi:10.1080/12265934.2011.580153
- Al Bavon, & Standerfer, C. (2010). The effect of the 1997 Texas motorcycle helmet law on motorcycle crash fatalities. *Southern Medical Journal*, 103(1), 11–17.
- Al-Haji, G. (2005). *Towards a road safety development index (RSDI), development of an international index to measure road safety performance*. Sweden: Linköping University.

- Allsop, R. E., Sze, N. N., & Wong, S. C. (2011). An update on the association between setting quantified road safety targets and road fatality reduction. *Accident Analysis and Prevention*, 43(3), 1279–83. doi:10.1016/j.aap.2011.01.010.
- Andrade, S. M., Soares, D. A., Matsuo, T., Liberatti, C. L. B., & Iwakura, M. L. H. (2008). Road injury-related mortality in a medium-sized Brazilian city after some preventive interventions. *Traffic Injury Prevention*, 9(5), 450–5. doi:10.1080/15389580802272831.
- Andreassen, D. (1991). Population and registered vehicle data vs. road deaths. *Accident Analysis and Prevention*, 23(5), 343–51.
- Andreassen, D. (1985). Linking deaths with vehicles and population. *Traffic Engineering & Control*, 26(11), 547–9.
- Anonymous (1835). *The highway act 1835*. The United Kingdom.
- Anonymous (2007). Mary Ward 1827–1869. Offaly Historical & Archaeological Society. Retrieved from <http://www.offalyhistory.com/reading-resources/history/famous-offaly-people/mary-ward-1827-69>.
- Anonymous (2010). Editorial safety science special issue road safety management. *Safety Science*, 48, 1081–4.
- Antić, B., Vujanić, M., Jovanović, D., & Pešić, D. (2011). Impact of the new road traffic safety law on the number of traffic casualties in Serbia. *Scientific Research and Essays*, 6(29), 6176–84. doi: 10.5897/SRE11.1419.
- Antoniou, C., & Yannis, G. (2013). State-space based analysis and forecasting of macroscopic road safety trends in Greece. *Accident Analysis and Prevention*, 60, 268–76. doi:10.1016/j.aap.2013.02.039.
- Antoniou, C., Yannis, G., & Papadimitriou, E. (2012). Macroscopic traffic safety data analysis and prediction. *Journal of transport and shipping*, 5, 97–116.
- Asch, P., Levy, D. T., Shea, D. & Bodenhorn, H. (1991). Risk compensation and the effectiveness of safety belt use laws: A case study of New Jersey. *Policy Sciences*, 24(2), 181–97.
- Assum, T., Bjørnskau, T., Fosser, S., & Sagberg, F. (1999). Risk compensation—the case of road lighting. *Accident Analysis and Prevention*, 31(5), 545–53.
- Bachani, A. M., Branching, C., Ear, C., Roehler, D. R., Parker, E. M., Tum, S., Hyder, A. A. (2013). Trends in prevalence, knowledge, attitudes, and practices of helmet use in Cambodia: results from a two year study. *Injury*, 44 Suppl 4, 31–7. doi:10.1016/S0020-1383(13)70210-9.

- Baguley, C., & M Shafii, M. (1995). Engineering approaches to reversing a worsening road accident trend in Malaysia. *International Forum on Road Safety Research*. Bangkok, Thailand, 25–27 October 1995.
- Baguley, C. (2001). The importance of a road accident data system and its utilisation. Paper Presented at *International Symposium on Traffic Safety Strengthening and Accident Prevention*. Nanjing, China, Nov 28-30. 2001.
- Balkin S., & Ord, J. K. (2001). Assessing the impact of speed limit increases on fatal interstate crashes. *Journal of Transportation and Statistics*, 4(1), 1–26.
- Ballart, X., & Riba, C. (1995). Impact of legislation requiring moped and motorbike riders to wear helmets. *Evaluation and Program Planning*, 18(4), 311–20.
- Baruya, A. (1998). *A review of speed-accident relationship for European roads* (PR/TT/029/97). United Kingdom: Transport Research Laboratory.
- Bédard, M., Stones, M. J., Guyatt, G. H., & Hirdes, J. P. (2001). Traffic-related fatalities among older drivers and passengers: past and future trends. *The Gerontologist*, 41(6), 751–6.
- Beenstock, M., & Gafni, D. (2000). Globalization in road safety: explaining the downward trend in road accident rates in a single country (Israel). *Accident Analysis and Prevention*, 32(1), 71–84.
- Bener, A. (2005). The neglected epidemic: road traffic accidents in a developing country, State of Qatar. *International Journal of Injury Control and Safety Promotion*, 12(1), 45–7. doi:10.1080/1745730051233142225.
- Bener, A. (2009). Emerging trend in motorisation and the epidemic of road traffic crashes in an economically growing country. *International Journal of Crashworthiness*, 14(2), 183–8. doi:10.1080/13588260802614399.
- Bergel-Hayat, R. (2012). Time-series models of aggregate road risk and their applications to European countries. *Transport Reviews*, 32(5), 653–70. doi:10.1080/01441647.2012.695407.
- Bester, C. J. (2001). Explaining national road fatalities. *Accident Analysis and Prevention*, 33(5), 663–72.
- Bijleveld, F. (1997). *Effectiveness of daytime motorcycle headlights in the European Union* (R-97-9). Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.
- Bijleveld, F., & Commandeur, J. (2006). *The basic evaluation model* (D-2006-2). Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.

- Bijleveld, F., Commandeur, J., Gould, P., & Koopman, S. J. (2008). Model-based measurement of latent risk in time series. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(1), 265–77.
- Bijleveld, F., Commandeur, J., Koopman, S. J., & Montfort, K. Van. (2010). Multivariate non-linear time series modelling of exposure and risk in road safety research. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 59(1), 145–61. doi:10.1111/j.1467-9876.2009.00690.x.
- Bjerver, K., & Goldberg, L. (1951). Alcohol and road traffic. *Proceedings of First International Conference on Alcohol and Traffic*. Stockholm.
- Bjørnskau, T. (2011). *Road traffic risk in Norway 2009–2010* (1164/2011). Oslo: Institute of Transport Economics.
- Bossche, F. Van Den, & Wets, G. (2003). *Macro models in traffic safety and the DRAG family: literature review*. Diepenbeek, Belgium: Universitaire campus.
- Bossche, F. Van Den, Wets, G., & Brijs, T. (2004). *A regression model with ARIMA errors to investigate the frequency and severity of road traffic accidents* (R-00-2002-01). Diepenbeek, Belgium: Universitaire campus.
- Bossche, F. Van Den, Wets, G., & Brijs, T. (2007). Analysis of road risk by age and gender category: time series approach. *Transportation Research Record*, 2019, 7–14. doi:10.3141/2019-02.
- Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *J. Amer. Statist. Ass.*, 70, 70–9.
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis forecasting and control* (Revised Ed.). Oakland, California: Holden-Day, Inc.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis forecasting and control* (3rd Ed.). Englewood Cliffs, New Jersey: Prentice Hall Inc.
- Box, G.E.P., & Jenkins, G.M. (1970). *Time series analysis, forecasting and control*. San Francisco: Holden-Day, Inc.
- British Medical Association (1960). *Relation of alcohol to road accidents*. London.
- Brocklebank, J. C., & David A. D. (2003). *SAS for forecasting time series* (2nd Ed.). Cary, NC: SAS Institute Inc.
- Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting* (2nd Ed.). New York: Springer-Verlag Inc.

- Broughton, J. (1988). Predictive models of road accident fatalities. *Traffic Eng. Control*, 29(5), 296-300.
- Broughton, J. (2009). *Post-2010 casualty forecasting*. United Kingdom: Transport Research Laboratory.
- Broughton, J., & Knowles, J. (2010). Providing the numerical context for British casualty reduction targets. *Safety Science*, 48(9), 1134–41. doi:10.1016/j.ssci.2010.01.008.
- Broughton, J., & Lawton, B. (2009). *Casualty forecasts for Northern Ireland in 2020 and Key Performance Indicators*. United Kingdom: Transport Research Laboratory.
- Brüde, U. (1995). What is happening to the number of fatalities in road accidents? A model for forecasts and continuous monitoring of development up to the year 2000. *Accident Analysis and Prevention*, 27(3), 405-10.
- Butterworth, A. K. (1932). Road accidents: the compensation problem. *Journal of Comparative Legislation and International Law*, 14(4), 189-200
- Byrne, S. J. (1920). Traffic regulation as a means of preventing injury and death from street vehicles. *American Journal of Public Health*, 10(4), 331-3.
- Campbell, B. J., Stewart, J. R., & Campbell, F. A. (1986). *Early results of seat belt legislation in the United State of America*. Chapel Hill: The University of North Carolina Highway Safety Research Center.
- Campos, V. R., de Souza e Silva, R., Duailibi, S., dos Santos, J. F., Laranjeira, R., & Pinsky, I. (2013). The effect of the new traffic law on drinking and driving in São Paulo, Brazil. *Accident Analysis and Prevention*, 50, 622–7. doi:10.1016/j.aap.2012.06.011.
- Carpenter, C. S., & Stehr, M. (2008). The effects of mandatory seatbelt laws on seatbelt use, motor vehicle fatalities, and crash-related injuries among youths. *Journal of Health Economics*, 27(3), 642–62. doi:10.1016/j.jhealeco.2007.09.010
- Carroll, P. S. (1973). Classifications of driving exposure and accident rates for highway safety analysis. *Accident Analysis and Prevention*, 5, 81–94.
- Cassie, A. B., & Allan W. R. (1961). Alcohol and road traffic accidents. *British Medical Journal*, 2(5268), 1668–71.
- CDC (2009). Leading causes of death, 1900–1998. Centers for Disease Control and Prevention. Retrieved from http://www.cdc.gov/nchs/nvss/mortality_historical_data.htm.

- Chalmers, N. (2010). Investigating fatal collisions. RoadPeace Conference 22 September, 2010 London. Retrieved from <http://www.roadpeace.org/>.
- Chapelon, J., & Lassarre, S. (2010). Road safety in France: The hard path toward science-based policy. *Safety Science*, 48(9), 1151–9. doi:10.1016/j.ssci.2010.04.015.
- Chatfield, C. (2000). *Time-series forecasting*. London: Chapman and Hall/CRC.
- Chi, G., Porter, J. R., Cosby, A. G., & Levinson, D. (2013). The impact of gasoline price changes on traffic safety: a time geography explanation. *Journal of Transport Geography*, 28, 1–11. doi:10.1016/j.jtrangeo.2012.08.015.
- Chu, B. Y. C., & Nunn, G. E. (1976). An analysis of the decline in California traffic fatalities during the energy crisis. *Accident Analysis & Prevention*, 8(3), 145–50. doi:10.1016/0001-4575(76)90016-6.
- Clark, D. E., & Cushing, B. M. (2004). Rural and urban traffic fatalities, vehicle miles, and population density. *Accident Analysis & Prevention* 36(6), 967–72. doi: 10.1016/j.aap.2003.10.006.
- Clyde, M. N. (1964). Michigan study indicates signals increase accidents. *Traffic Engineering*, 35(2), 32.
- Cobb, R. W., & Coughlin, J. F. (1998). Are elderly drivers a road hazard?: Problem definition and political impact. *Journal of Aging Studies*, 12(4), 411–27. doi:10.1016/S0890-4065(98)90027-5.
- Cohen, A., & Einav, L. (2003). The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *The Review of Economics and Statistics*, 85(4), 828–43.
- Cohen, J., Dearnaley, E. J., & Hansel, C. E. M. (1958). The risk taken in driving under the influence of alcohol. *British Medical Journal*, 1(5085): 1438–42.
- Colton, T., & Buxbaum, R. C. (1968). Motor vehicle inspection and motor vehicle accident mortality. *AJPH*, 58(6), 1090–9.
- Commandeur, J. J. F., Bijleveld, F. D., Bergel-Hayat, R., Antoniou, C., Yannis, G., & Papadimitriou, E. (2013). On statistical inference in time series analysis of the evolution of road safety. *Accident Analysis and Prevention*, 60, 424–34. doi:10.1016/j.aap.2012.11.006
- Conybeare, J. A. C. (1980). Evaluation of automobile safety regulations: the case of compulsory seat belt legislation in Australia. *Policy Sciences*, 12(1), 27–39. doi:10.1007/BF00137361.

- Corben, B. F., Logan, D. B., Fanciulli, L., Farley, R., & Cameron, I. (2010). Strengthening road safety strategy development "Towards Zero" 2008–2020 – Western Australia's experience scientific research on road safety management, SWOV workshop 16 and 17 November 2009. *Safety Science*, 48(9), 1085–97. doi:10.1016/j.ssci.2009.10.005.
- Council, F M., & Stewart, J. R. (1999). Safety effects of the conversion of rural two-lane to four-lane roadways based on cross-sectional models. Paper presented at *Transportation Research Board 78th Annual Meeting*. Washington DC.
- Dalvi, Q. (2004). Review Paper. *Transport Reviews*, 24(3), 365–76. doi:10.1080/0144164042000231404.
- De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22(3), 443–73. doi:10.1016/j.ijforecast.2006.01.001.
- De Pauw, E., Daniels, S., Thierie, M., & Brijs, T. (2014). Safety effects of reducing the speed limit from 90 km/h to 70 km/h. *Accident Analysis and Prevention*, 62(2014), 426–31.
- Dee, T. S. (2009). Motorcycle helmets and traffic safety. *Journal of Health Economics*, 28(2), 398–412. doi:10.1016/j.jhealeco.2008.12.002.
- Department of Statistics (2012). *Population projections Malaysia 2010–2040*. Kuala Lumpur: Prime Minister Office.
- Deutermann, W. V. (2005). *Calculating lives saved by motorcycle helmets* (DOT-HS-809-861). Washington, DC: National Highway Traffic Safety Administration.
- DfT (2009). *Reported road casualties Great Britain: 2009 annual Report*. London.
- EC (2004). *COST action 329. Models for traffic and safety development and interventions*. Luxembourg.
- ECMT (1984). *Statistical report on road accidents in 1981*. Paris.
- Elvik, R. (1993). Quantified road safety targets: a useful tool for policy making? *Accident Analysis and Prevention*, 25(5), 569–83.
- Elvik, R. (2010). Sources of uncertainty in estimated benefits of road safety programmes. *Accident Analysis and Prevention*, 42(6), 2171–8. doi:10.1016/j.aap.2010.03.022.
- Elvik, R. (2010a). The stability of long-term trends in the number of traffic fatalities in a sample of highly motorised countries. *Accident Analysis and Prevention*, 42(1), 245–60. doi:10.1016/j.aap.2009.08.002.

- Elvik, R., Vaa, T., Erke, A., & Sorensen, M. (2009). *Handbook of road safety measures* (2th Ed.). The United Kingdom: Emerald Group Publishing.
- Emenalo, S., Puustelli, M., Ciampi, A., & Joshi, H. P. (1977). Analysis of road traffic accidents data in Zambia. *Accident Analysis and Prevention*, 9, 81–91.
- Enders, W. (2003). *Applied econometric time series*. New York: John Wiley & Sons Inc.
- Evans, L. & Frick, M. C. (1988). Helmet effectiveness in preventing motorcycle driver and passenger fatalities. *Accident Analysis and Prevention*, 20, 447–58.
- Evans, L. (1986). The effectiveness of safety belts in preventing fatalities. *Accident Analysis and Prevention*, 18(3), 229–41.
- Evans, L. (2004). *Traffic safety*. Bloomfield Hills, MI: Science Serving Society.
- Evans, W. N., & Graham, J. D. (1991). Risk reduction or risk compensation? The case of mandatory safety-belt use laws. *Journal of Risk and Uncertainty* January, 4(1), 61–73.
- Farmer, C. M. (1997). Trends in motor vehicle fatalities. *Journal of Safety Research*, 28(1), 37–48. doi:10.1016/S0022-4375(96)00032-1.
- Farmer, C. M., Retting, R. A., & Lund, A. K. (1999). Changes in motor vehicle occupant fatalities after repeal of the national maximum speed limit. *Accident Analysis and Prevention*, 31, 537–43.
- Feng, C. (2001). Synthesis of studies on speed and safety. *Transportation Research Record*, 1779(01), 86–92.
- Fergusson D. M., Swain-Campbell N. R., & Horwood L. J. (2003). Risky driving behaviour in young people: Prevalence, personal characteristics and traffic accidents. *Australian & NZ Journal of Public Health*, 27(3): 337–42.
- Fernandes, F. A. O., & Alves de Sousa, R. J. (2013). Motorcycle helmets—a state of the art review. *Accident Analysis and Prevention*, 56, 1–21. doi:10.1016/j.aap.2013.03.011.
- Fleming, N. S., & Becker, E. R. (1992). The impact of the Texas 1989 motorcycle helmet law on total and head-related fatalities, severe injuries, and overall injuries. *Medical Care*, 30(9), 832–45.
- Freeman, C. P. (1985). Isolated pancreatic damage following seat belt injury. *Injury*, 16(7), 478–80.
- Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., & Thomsen, L. K. (1995). Measuring the contribution of randomness, exposure, weather, and daylight to

- the variation in road accident counts. *Accident Analysis and Prevention*, 27(1), 1–20.
- Friedman, L. S., Barach, P., & Richter, E. D. (2007). Raised speed limits, case fatality and road deaths: a six year follow-up using ARIMA models. *Injury Prevention*, 13, 156–61. doi:10.1136/ip.2006.014027.
- Garbacz, C. (1989). Traffic fatalities in Taiwan. *Journal of Transport Economics and Policy*, 23(3), 317–27.
- Garber, S., & Graham, J. D. (1990). The effects of the new 65 mile-per-hour speed limit on rural highway fatalities: a state-by-state analysis. *Accident Analysis and Prevention*, 22(2), 137–49.
- Gargett, S., Connelly, L. B., & Nghiem, S. (2011). Are we there yet? Australian road safety targets and road traffic crash fatalities. *BMC Public Health*, 11(1), 270. doi:10.1186/1471-2458-11-270.
- Garrett, J. W., & Braunstein, P. W. (1962). The seat belt syndrome. *Journal of Trauma Injury Infection & Critical Care*, 2(3), 220-38.
- Gaudry, M. (2002). *DRAG, model of the demand for road use, accidents and their severity, applied in Quebec from 1956 to 1982*. Technical Report Publication. Montreal: Département de sciences économiques, Université de Montréal.
- Gaudry, M., & de Lapparent, M. (2013a). Part 1. National road safety performance: Data, the emergence of two single-outcome modeling streams and public health. *Research in Transportation Economics*, 37(1), 6–19. doi:10.1016/j.retrec.2012.02.001
- Gaudry, M., & de Lapparent, M. (2013b). Part 3. Multivariate road safety models: Future research orientations and current use to forecast performance. *Research in Transportation Economics*, 37(1), 38–56. doi:10.1016/j.retrec.2012.02.003
- Gaudry, M., & Himouri, S. (2013). *DRAG-ALZ-1*, a first model of monthly total road demand, accident frequency, severity and victims by category, and of mean speed on highways, Algeria 1970–2007. *Research in Transportation Economics*, 37(1), 66–78. doi:10.1016/j.retrec.2011.08.010.
- Gaudry, M., & Lassarre, S. (2000). *The international DRAG family*. Oxford: Elsevier Science.
- Gerdtham, U., & Ruhm, C. J. (2002). *Deaths rise in good economic times: evidence from the OECD* (9357). Cambridge. Retrieved from <http://www.nber.org/papers/w9357>.

- Gharaybeh, F. A. (1994). Application of Smeed's formula to assess development of traffic safety in Jordan. *Accident Analysis and Prevention*, 26(1), 113–20.
- Glanville, W. H. (1951). Road safety and road research. *Journal of the Royal Society of Arts*, 99(4838), 144–92.
- Glanville, W. H. (1954). Safety on the road. *Journal of the Royal Society of Arts*, 102(4926), 496–519.
- Gossner, O., & Picard, P. (2005). On the consequences of behavioral adaptations in the cost–benefit analysis of road safety measures. *The Journal of Risk and Insurance*, 72(4), 577–99.
- Graham, J. D. (1982). On Wilde's theory of risk homeostasis. *Risk analysis*, 2, 235–7.
- Graham, J. D., & Lee, Y. (1986). Behavioral response to safety regulation: the case of motorcycle helmet-wearing legislation. *Policy Sciences*, 19(3), 253–73.
- Haddon, W. (1968). The changing approach to the epidemiology, prevention, and amelioration of trauma: The transition to approaches etiologically rather than descriptively based. *AJPH*, 58(8), 1431–8.
- Haight, F. A. (1984). Why the per capita traffic fatality rate is falling. *Journal of Safety Research*, 15(4), 137–40. doi:10.1016/0022-4375(84)90044-6.
- Haight, F. A. (1994). Problems in estimating comparative costs of safety and mobility. *Journal of Transport Economics and Policy*, 28(1), 7–30.
- Hakim, S. (1991). A critical review of macro models for road accidents. *Accident Analysis and Prevention*, 23(5), 379–400.
- Hakkert, A. S., Gitelman, V., & Vis, M. A. (2007). *Road Safety Performance Indicators: Theory*. Deliverable D3.6 of the EU FP6 Project SafetyNet.
- Hakkert, S., & Braimaister, L. (2002). *The uses of exposure and risk in road safety studies*. Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.
- Halsey, M. (1941). *Traffic accidents and congestion*. New York: John Wiley & Sons Inc.
- Hanke, J. E., & Wichern, D. W. (2005). *Business forecasting* (8th Ed.). Upper Saddle River, New Jersey: Pearson Education Inc.
- Haque, M. O., & Cameron, M. (1989). Effect of the Victorian zero BAC legislation on serious casualty accidents: July 1984–December 1985. *Journal of Safety Research*, 20(3), 129–37. doi:10.1016/0022-4375(89)90058-3.

- Harvey, A.C., & Durbin, J. (1986). The effects of seat belt legislation on British road casualties: a case study in structural time series modeling. *J. R. Statist. Soc. A*, 149(3), 187–227.
- Hauer, E. (2010). On prediction in road safety. *Safety Science*, 48(9), 1111–22. doi:10.1016/j.ssci.2010.03.003.
- Hayakawa, H., Fischbeck, P. S., & Fischhoff, B. (2000). Traffic accident statistics and risk perceptions in Japan and the United States. *Accident Analysis and Prevention*, 32(6), 827–35.
- Hazen, A., & Ehiri, J. E. (2006). Road traffic injuries: hidden epidemic in less developed countries. *Journal of the National Medical Association*, 98(1), 73–82.
- He, J., King, M., Watson, B., Rakotonirainy, A., & Fleiter, J. (2013). Speed enforcement in China: National, provincial and city initiatives and their success. *Accident Analysis and Prevention*, 50, 282–8. doi:10.1016/j.aap.2012.04.017
- Hedlund, J. (2000). Risky business: safety regulations, risk compensation, and individual behavior. *Injury Prevention*, 6(2), 82–9. doi:10.1136/ip.6.2.82.
- Hedlund, J., Arnold, R., Cerrelli, E., Hoxie, P., & Skinner, D. (1984). An assessment of the 1982 traffic fatality decrease. *Accident Analysis and Prevention*, 16(4), 274–61.
- Hedlund, J., Shults, R. A., & Compton, R. (2003). What we know, what we don't know, and what we need to know about graduated driver licensing. *Journal of Safety Research*, 34(1), 107–15.
- Hermans, E., Wets, G., & Bossche, F. Van Den (2006). *Frequency and severity of Belgian road traffic accidents studied by state space methods*. Diepenbeek, Belgium: Hasselt University Transportation Research Institute.
- Hermans, E., Wets, G., & Bossche, F. Van Den. (2005). *Describing the evolution in the number of highway deaths by a decomposition in exposure, accident risk and fatal risk*. Diepenbeek, Belgium: Hasselt University Transportation Research Institute.
- Hess, S., & Polak, J. (2003). *An analysis of the effects of speed limit enforcement cameras on accident rates*, 1–16. London: Centre for Transport Studies Imperial College.
- Hilton, M. (1984). The impact of recent changes in California drinking-driving laws on fatal accident levels during the first post intervention year: An interrupted time series analysis. *Law and Society Review*, 18(4), 605–27.

- Hoskin, A. E. (1986). Consequences of raising the speed limit. *Journal of Safety Research*, 17, 179-82.
- Houston, D. J., & Richardson, L. E. (2002). Traffic safety and the switch to a primary seat belt law: the California experience. *Accident Analysis and Prevention*, 34(6), 743–51.
- Huelke, D. F., & Gikas, P. W. (1968). Causes of deaths in automobile accidents. *JAMA*, 203(13): 1100-7. doi:10.1001/jama.1968.03140130012003.
- Hurt, H. H., Ouellet, J. V., Thom, D. R. (1981). *Motorcycle accident cause factors and identification of countermeasures: Technical report* (Vol. I). Washington, DC: National Highway Traffic Safety Administration.
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(2006): 679 – 688. doi:10.1016/j.ijforecast.2006.03.001
- Ipingbemi, O. (2008). Spatial analysis and socio-economic burden of road crashes in south-western Nigeria. *International Journal of Injury Control and Safety Promotion*, 15(2), 99–108. doi:10.1080/17457300802150785.
- Izquierdo, F. A., Ramírez, B. A., McWilliams, J. M. M., & Ayuso, J. P. (2011). The endurance of the effects of the penalty point system in Spain three years after. Main influencing factors. *Accident Analysis and Prevention*, 43(3), 911–22. doi:10.1016/j.aap.2010.11.014.
- Izquierdo, F. A., Ramírez, B. A., & Rodríguez, E. B. (2013). The interurban DRAG-Spain model: The main factors of influence on road accidents in Spain. *Research in Transportation Economics*, 37(2013) 57–65.
- Jacobs, G., Aeron-Thomas, A., & Astrop, A. (2000). *Estimating global road fatalities* (445). United Kingdom: Transport Research Laboratory.
- Jacobs, G.D. (1982). The potential for road accident reduction in developing countries. *Transport Reviews*, 2(2), 213–24.
- Jacobs, G.D., & Cutting, C.A. (1986). Further research in accident rates in developing countries. *Accident Analysis and Prevention*, 18(2), 119–27.
- Jacobs, G.D., & Hutchinson, P. (1973). *A study of accident rates in developing countries* (546). Crowthorne, Berkshire: Transport and Road Research Laboratory.
- Jaeger, L., & Lassarre, S. (2000). *The TAG-1 model for France. In: structural road accident models: The international DRAG family* (Gaudry, M., & Lassarre, S. Eds.). Oxford: Elsevier Science.

- Jamilah, M. M., Norlen M., Mohd Rasid O., Fuad A., Mohd Faudzi M. Y., & Wong, S. V. (Eds.). (2012), *Evaluation of the effectiveness of Ops Bersepadu Hari Raya 2011: Conducted over the Hari Raya Period from 23 August 2011 to 6 September 2011* (MRR 03/2012). Kuala Lumpur: Malaysian Institute of Road Safety Research.
- Janoff, M. S., & Cassel, A. (1971). Effect of daytime motorcycle headlight laws on motorcycle accidents. *Highway Research Record*, 377, 53–63.
- Jessie, W. S., & Yuan, W. (1998). The efficacy of safety policies on traffic fatalities in Singapore. *Accident Analysis and Prevention*, 30(6), 745–54.
- Johansson, P. (1996). Speed limitation and motorway casualties: a time series count data regression approach. *Accident Analysis and Prevention*, 28(1), 73–87.
- Joks, H. C. (1976). The effects of automobile safety regulation: comments on Peltzman's reply. *Accident Analysis and Prevention*, 8(3), 213–4.
- Joks, H. C., & Wuerdemann H. (1970). *Estimating the effects of crash phase injury countermeasures* (4013-395). Springfield, Virginia: The Travelers Research Corporation, National Technical Information Service.
- Jones, A.P., Haynes, R., Kennedy, V., Harvey, I. M., Jewell, T., & Lea, D. (2008). Geographical variations in mortality and morbidity from road traffic accidents in England and Wales. *Health & Place*, 14(3): 519–35. doi: 10.1016/j.healthplace. 2007.10.001.
- Jorgensen, N. O. (1969). A model for forecasting traffic accidents. *Accident Analysis and Prevention*, 1, 261–6.
- Kamarudin, A., Rozmi, I., Riza, A. A., & Muhamad Nazri, B. (2010). Prediction of helmet use among Malaysian motorcyclist using structural equation modeling. *Australian Journal of Basic and Applied Sciences*, 4(10), 5263–70.
- Kanchan, T., Kulkarni, V., Bakkannavar, S. M., Kumar, N., & Unnikrishnan, B. (2012). Analysis of fatal road traffic accidents in a coastal township of South India. *Journal of Forensic and Legal Medicine*, 19(8), 448–51. doi:10.1016/j.jflm.2012.02.031.
- Kanellaidis, G., & Vardaki, S. (2011). Highway geometric design from the perspective of recent safety developments. *Journal of Transportation Engineering*, 137(12), 841–4. doi:10.1061/(ASCE)TE.1943-5436.0000322.
- Karlaftis, M. G., & Tarko, A. P. (1998). Heterogeneity considerations in accident modeling. *Accident Analysis and Prevention*, 30(4), 425–33.

- Kim, K. S., Myeong, M. H., & Kweon, Y. (2006). Evaluating the effects of safety policy measures on traffic fatalities in Korea. *Transport Reviews*, 26(3), 293–304. doi:10.1080/01441640500424088.
- Kim, K., Li, L., Richardson, J., & Nitz, L. (1998). Drivers at fault : influences of age, sex, and vehicle type. *Journal of Safety Research*, 29(3), 171–9.
- Kloeden C.N., McLean A.J., & Glonek, G. (2002). *Reanalysis of travelling speed and the risk of crash involvement in Adelaide South Australia* (CR 207). Australian Transport Safety Bureau.
- Koornstra, M. J. (1988). Development of road safety in some European countries and the USA. A theoretical and quantitative mathematical analysis. Paper presented to the *Conference Road Safety in Europe*. Gothenburg, Sweden, 12-14 October, 1988.
- Koornstra, M. J. (2007). Prediction of traffic fatalities and prospects for mobility becoming sustainable-safe. *Sadhana*, 32(4), 365–95.
- Kopits, E., & Cropper, M. (2003). *Traffic fatalities and economic growth* (3035). Washington DC: The World Bank.
- Kopits, E., & Cropper, M. (2005). Traffic fatalities and economic growth. *Accident Analysis and Prevention*, 37(1), 169–78. doi:10.1016/j.aap.2004.04.006.
- Koren, C., & Borsos, A. (2011). Is Smeed's Law still valid a world-wide analysis of the trends in fatality rates, *Journal of Society for Transportation and Traffic Studies*, 1, 64–76.
- Krystek, R., & Żukowska, J. (2007). Time series – the tool for traffic safety analysis. *Zagadnienia Eksploatacji Maszyn*, 1(149), 173–82.
- Kulanthayan, S., Radin Umar, R. S., Hariza, H.A. Mohd Nasir, M. T., & Harwant, S. (2000). Compliance of proper safety helmet usage in motorcyclists. *Medical Journal of Malaysia*, 55(1), 40–4.
- Kweon, Y., & Kockelman, K. M. (2004). Spatially disaggregate panel models of crash and injury counts: the effect of speed limits and design. Paper presented at *Transportation Research Board 83rd Annual Meeting*. Washington DC.
- Lassarre, S. (2001). Analysis of progress in road safety in ten European countries. *Accident Analysis and Prevention*, 33(6), 743–51.
- Lateef, M. U. (2010). Estimation of fatalities due to road traffic crashes in Karachi, Pakistan, using capture-recapture method. *Asia-Pacific Journal of Public Health / Asia-Pacific Academic Consortium for Public Health*, 22(3), 332–41. doi:10.1177/ 1010539509356808.

- Lateef, M. U. (2011). Spatial patterns monitoring of road traffic injuries in Karachi metropolis. *International Journal of Injury Control and Safety Promotion*, 18(2), 97–105. doi:10.1080/17457300.2010.490588.
- Lave, C. A. (1985). Speeding coordination and the 55-mph limit. *American Economic Review*, 75, 1159-64.
- Lave, C., & Elias, P. (1994). Did the 65 mph speed limit save lives? *Accident Analysis and Prevention*, 26(1), 49–62.
- Law, T. H., Umar, R. S. R., Zulkaurnain, S., & Kulanthayan, S. (2005). Impact of the effect of economic crisis and the targeted motorcycle safety programme on motorcycle-related accidents, injuries and fatalities in Malaysia. *International Journal of Injury Control and Safety Promotion*, 12(1), 9–21. doi:10.1080/17457300512331339166.
- Ledolter, J., & Chan, K. S. (1996). Evaluating the impact of the 65 mph maximum speed limit on Iowa rural interstates. *The American Statistician*, 50(1), 79–85.
- Letirand, F., & Delhomme, P. (2005). Speed behaviour as a choice between observing and exceeding the speed limit. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(6), 481–92. doi:10.1016/j.trf.2005.06.002.
- Lin, M., & Kraus, J. (2008). Methodological issues in motorcycle injury epidemiology. *Accident Analysis and Prevention*, 40(5), 1653–60.
- Liu B. C., Ivers R., Norton R., Boufous S., Blows S., Lo S. K. (2008). Helmets for preventing injury in motorcycle riders. *Cochrane Database of Systematic Reviews*, 2008(1). doi: 10.1002/14651858.CD004333.pub3.
- Loo, B. P. Y., Hung, W. T., Lo, H. K., & Wong, S. C. (2005). Road safety strategies: a comparative framework and case studies. *Transport Reviews*, 25(5), 613–39. doi:10.1080/01441640500115892.
- Lu, M. (2007). Analysis and forecast of road traffic safety improvement in the Netherlands. Paper presented at *Transportation Research Board 87th Annual Meeting*. Washington DC.
- Lund H.V. (1979). *Effect of motor cyclists' use of running lights based on accident data during the period of 1st May, 1977 – 30th June, 1978*. Lyngby, Denmark: Danish Council of Road Safety Research.
- Luoma, J., & Sivak, M. (2013). *Why is road safety in the U.S. not on par with Sweden, the U.K., and the Netherlands? Lessons to be learned* (UMTRI-2013-1). Michigan: Transportation Research Institute, the University of Michigan.

- Mackay, G. M. (1969). Some features of traffic accidents. *British Medical Journal*, 4(5686), 799–801.
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 16(4), 451–76. doi:10.1016/S0169-2070(00)00057-1.
- Males, M. (2007). California's graduated driver license law: effect on teenage drivers' deaths through 2005. *Journal of Safety Research*, 38(6), 651–9. doi:10.1016/j.jsr.2007.09.002.
- Malnaca, K. (2008). Risk homeostasis theory in traffic safety. In *21st ICTCT workshop*.
- Malo, A. F. (1967). *Signal modernisation* (93). Washington DC: HRB.
- Mayhew, D. R., Simpson, H. M., & Ferguson, A. F. (2006). Collisions involving senior drivers: high-risk conditions and locations. *Traffic Injury Prevention*, 7(2), 117–24. doi:10.1080/15389580600636724.
- Mayrose, J. (2008). The effects of a mandatory motorcycle helmet law on helmet use and injury patterns among motorcyclist fatalities. *Journal of Safety Research*, 39, 429–32. doi:10.1016/j.jsr.2008.07.001.
- McCarthy, P. S. (1999). Public policy and highway safety: a city-wide perspective. *Regional Science and Urban Economics*, 29(2), 231–44.
- McFarland, R. A., Roland C. M., & Warren, A. B. (1955). *Human variables in motor vehicle accidents*. Boston: School of Public Health, Harvard University.
- McKnight, A. J., & Klein, T. M. (1990). Relationship of 65-mph limit to speeds and fatal accidents. *Transportation Research Record*, 1281, 71–7.
- McManus, W. (2007). *The economics of road safety: an international perspective* (UMTRI-2007-23). Michigan: Transportation Research Institute, the University of Michigan.
- Meier, K. J., & Morgan, D. R. (1980). Speed kills: a longitudinal analysis of traffic fatalities and the 55 mph speed limit. Paper presented at the *Southwestern Political Science Association annual meeting*, April 2-5, 1980, Houston, Texas.
- Mekky, A. (1985). Effects of rapid increase in motorization levels on road fatality rates in some rich developing countries. *Accident Analysis and Prevention*, 17(2), 101–9.
- Milton, J., & Mannering, F. (1996). *The relationship between highway geometrics, traffic related elements and motor vehicle*. Washington State Dept. of Transportation.

- Milton, J., & Mannering, F., (1998). *The relationship among highway geometrics, traffic related elements and motor vehicle accident frequencies*. *Transportation*, 25(4), 395–413.
- Minter, A. L. (1987). Road casualties improvement by learning processes. *Traffic Engineering & Control*, 28, 74-9.
- MIROS (2014). Revision on Helmet Standard MS1:1996 to MS1:2011. Retrieved from <http://www.miros.gov.my/web/>.
- Mohammadi, G. (2009). Mobile phone and seat belt usage and its impact on road accident fatalities and injuries in southeast Iran. *International Journal of Crashworthiness*, 14(4), 309–14. doi:10.1080/13588260802671423.
- Mohan, D. (2002). Road safety in less-motorized environments : *International Journal of Epidemiology*, (31), 527–32.
- Mohan, D. (2004). *The road ahead: traffic injuries and fatalities in India*. New Delhi: Indian Institute of Technology Delhi.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2008). *Introduction to time series analysis and forecasting*. Hoboken, New Jersey: John Wiley & Sons Inc.
- Mountain, L. J., Hirst, W. M., & Maher, M. J. (2005). Are speed enforcement cameras more effective than other speed management measures? The impact of speed management schemes on 30 mph roads. *Accident Analysis and Prevention*, 37(4), 742–54. doi:10.1016/j.aap.2005.03.017
- Muhammad Marizwan, A. M., Jonsson, T., & Varhelyi, A. (2013). Development of a safety performance function for motorcycle accident fatalities on Malaysian primary roads. *Safety Science*, 60, 13–20.
- Muller, A. (1982). An evaluation of the effectiveness of motorcycle daytime headlight laws. *Am J Public Health* 1982, 72(10), 1136–41.
- Muller, A. (1985). Motorcycle headlight-use laws: a contrasting view. *Am J Public Health* 1985, 75(5), 547–49.
- Muller, A. (2004). Florida's motorcycle helmet law repeal and fatality rates. *American Journal of Public Health*, 94(4), 556–8. doi:10.2105/AJPH.94.4.556
- Murry, J. P., Stam, A., & Lastovicka, J. L. (1993). Evaluating an anti-drinking campaign with a sample survey and time series intervention analysis. *Journal of the American Statistical Association*, 88(421), 50–6.
- Nagata, T., Setoguchi, S., Hemenway, D., & Perry, M. J. (2008). Effectiveness of a law to reduce alcohol-impaired driving in Japan. *Injury Prevention : Journal of the*

International Society for Child and Adolescent Injury Prevention, 14(1), 19–23.
doi:10.1136/ip.2007.015719.

- Neustrom, M. W., & Norton, W. M. (1993). The impact of drunk driving legislation in Louisiana. *Journal of Safety Research*, 24(2), 107–21. doi:10.1016/0022-4375(93)90005-8.
- NHTSA (1984). *Final regulation impact analyses, amendment to federal motor vehicle safety standard 208, passenger car front seat occupant protection* (DOT HS-806-572). Washington DC: National Highway Traffic Safety Administration.
- NHTSA (1996). *Effectiveness of occupant protection systems and their use*. Third Report to Congress. National Highway Traffic Safety Administration.
- NHTSA (1998a). *Further analysis of motorcycle helmet effectiveness using codes linked data*. National Center for Statistics and Analysis Research Note. Washington, DC: National Highway Traffic Safety Administration.
- NHTSA (1998b). *The effect of increased speed limits in the Post-NMSL era*. Washington DC: Department of Transport.
- Nilsson, G. (2004). *Traffic safety dimensions and the power model to describe the effect of speed on safety* (Unpublished doctoral dissertation). Lund, Sweden: Lund Institute of Technology.
- Noland, R. B. (2003). Medical treatment and traffic fatality reductions in industrialized countries. *Accident Analysis and Prevention*, 35(2003), 877–83.
- Noland, R. B. (2004). Motor vehicle fuel efficiency and traffic fatalities. *Energy Journal*, 25(4), 1–22.
- Noland, R.B. (2001). Traffic fatalities and injuries: are reductions the result of 'improvements' in highway design standards. Paper presented at *Transportation Research Board 80th Annual Meeting*. Washington DC: Transportation Research Board.
- Noland, R.B., & Oh, L. (2004). The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. *Accident Analysis and Prevention*, 36, 525–32.
- Norlen M., Wahida A. B., Nurul Kharmila A., Muhammad Fadhli M. Y., Sharifah Allyana S. M. R., & Mohd Shah M. (2012). *Alcohol and drug use among fatally injured drivers in urban area of Kuala Lumpur* (MRR 02/2012). Kuala Lumpur: Malaysian Institute of Road Safety Research.
- Norman, L. G. (1962). *Road traffic accidents, epidemiology, control and prevention*. Public Health Paper 12. Geneva: World Health Organization.

- O'Cinneide, D., & Troutbeck, R. J. (1995). At-grade intersections: Worldwide review, in *Proc of the International Symposium on Highway Geometric Design Practices*. Transportation Research Board. August 30–September 1, 1995. Boston, Massachusetts.
- OECD (1990). *Behavioral adaptations to changes in the road transport system*. University of Chicago Press.
- ONISR (2001). Les grandes donne'es de l'accidentologie.
- Oppe, S. (1989). Macroscopic models for traffic and traffic safety. *Accident Analysis and Prevention*, 21(3), 225–232.
- Oppe, S. (2001). *Traffic safety developments in Poland* (D-2001-8). Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.
- Oster, C. V., & Strong, J. S. (2013). Analyzing road safety in the United States. *Research in Transportation Economics*, 43(1), 98–111. doi:10.1016/j.retrec. 2012.12.005.
- Ouellet, J. V. (2011). Helmet use and risk compensation in motorcycle accidents. *Traffic Injury Prevention*, 12(1), 71–81. doi:10.1080/15389588.2010.529974.
- Page, Y. (2001). A statistical model to compare road mortality in OECD countries. *Accident Analysis and Prevention*, 33(3), 371–85.
- Pankratz, A. (1991). *Forecasting with dynamic regression models*. New York: John Wiley & Sons Inc.
- Partyka, S. C. (1984). Simple models of fatality trends using employment and population data. *Accident Analysis and Prevention*, 16(3), 211–22.
- PDRM (1995). *Annual road accident statistics report 1994*. Kuala Lumpur: Royal Malaysian Police.
- PDRM (2013). *Annual road accident statistics report 1969–2012*. Kuala Lumpur: Royal Malaysian Police.
- Peltzman, S. (1975). The effects of automobile safety regulation. *Journal of Political Economy*, 83(4), 677–726.
- Perlot, A., & Prower, S. (2003). *Review of the evidence for motorcycle and motorcar daytime lights*. Bruxelles, Belgium: Federation of European Motorcyclists' Associations.
- Pisano, P. A., Goodwin, L. C., & Rossetti, M. A. (2008). *U.S. highway crashes in adverse road weather conditions*. Washington DC: U.S. DOT Federal Highway Administration.

- Pramada Valli, P. (2005), Road accident models for large metropolitan cities of India, *IATSS Research*, 29(1), 57–65.
- Preusser, D.F., Lund, A.K., Williams, A.F., & Blomberg, R.D. (1988). Belt use by high risk drivers before and after New York's seat belt use law. *Accident Analysis and Prevention*, 20(4), 245–50.
- Rabihah, I, Azli, A.Z., Abdullah, S. & Nurulhana, B. (2013). Enforcement and helmet use compliance: an observation in selected areas in selangor. Paper presented at *4th International Conference for Road Safety*, 4-6 March 2013, Sydney, Australia.
- Radin Umar, R. S. (1998). Fatality models for Malaysia: towards year 2000. *Pertanika Journal Science and Technology*, 6(2), 1–13.
- Radin Umar, R. S. (2005). Updates of road safety status in Malaysia. *IATSS Research*, 29(1), 78–80.
- Radin Umar, R. S. (2007). Integrated approach of road safety in malaysia. *The 7th Malaysian Road Conference*. Sunway, Malaysia.
- Radin Umar, R. S., Mackay, M., & Hills, B. (2000). Multivariate analysis of motorcycle accidents and the effects of exclusive motorcycle lanes in Malaysia. *Journal of Crash Prevention and Injury Control*, 2(1), 11–7.
- Radin Umar, R., Mackay, M. G., & Hills, B. L. (1996). Modelling of conspicuity-related motorcycle accidents in Seremban and Shah Alam, Malaysia. *Accident Analysis and Prevention*, 28(3), 325–32.
- Raftery, S. J., Grigo, J. A. L., & Woolley, J. E. (2011). Heavy vehicle road safety: a scan of recent literature. *Journal of the Australasian College of Road Safety*, 22(3), 18–24.
- Ray, C. (1989). Legal control of drunken driving : a time series study of California data. *International Journal of Forecasting*, 5(1989), 515–22.
- Reurings, M., & Commandeur, J. (2007). *International orientation on methodologies for modelling developments in road safety*. Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.
- Reynolds, D. J. (1956). The cost of road accidents. *J. Royal Stat. Soc. A*, 119, 393–408.
- Richardson, L. E., & Shaw, K. B. (2009). Traffic safety policy in the European Union: the role of interest groups. Paper prepared for presentation at the *biennial meetings of the European Union Studies Association*. Los Angeles, California, April 23-25, 2009.

- Rizzi, L. I., Cumsille, S., Fresard, F., Gazmuri, P., & Muñoz, J. C. (2011). Cost-effective measures for reducing road fatalities in the short term. *Transport Reviews*, 31(1), 1–24. doi:10.1080/01441641003736572.
- Robertson, L. S. (1977). State and federal new-car safety regulation: effects on fatality rates. *Accident Analysis and Prevention*, 9(3), 151–6. doi:10.1016/0001-4575(77) 90016-1.
- Robertson, L. S. (1981). Automobile safety regulations and death reductions in the United States. *American Journal of Public Health*, 71(8), 818–22.
- Roh, J. W., Bessler, D. a, & Gilbert, R. F. (1999). Traffic fatalities, Peltzman's model, and directed graphs. *Accident Analysis and Prevention*, 31(1-2), 55–61.
- Rohayu S., Sharifah Allyana S. M. R, Jamilah M. M., & Wong, S. V. (2012). *Predicting Malaysian road fatalities for year 2020* (MRR 06/2012). Kuala Lumpur: Malaysian Institute of Road Safety Research.
- Ross, H. L. (1973). Law, science, and accidents: the British road safety act of 1967. *The Journal of Legal Studies*, 2(1), 1–78.
- Ross, H. L., & Klette, H. (1995). Abandonment of mandatory jail for impaired drivers in Norway and Sweden. *Accident Analysis and Prevention*, 27(2), 151–7.
- Roth, H., & Lombard, C. (1953). *Crash helmet*. US patent No. 2,625,683. Washington, DC: U.S. Patent and Trademark Office.
- Routley, V., Ozanne-Smith, J., Li, D., Yu, M., Wang, J., Wu, M., Qin, Y. (2009). Attitudes to seat belt wearing and related safety features in two cities in China. *International Journal of Injury Control and Safety Promotion*, 16(1), 15–26. doi:10.1080/17457300902738752.
- Sari, M., Mutlu, Ö., & Zeytinoğlu, A. (2009). Effects of human and external factors on traffic accidents. *Matematica - Informatica- Fizica*, LXI(1), 9–18.
- SAS Institute Inc. (2004). *SAS/ETS 9.1 User's Guide*. Cary, NC: SAS Institute Inc.
- Sauerzapf, V., Jones, A. P., & Haynes, R. (2010). The problems in determining international road mortality. *Accident Analysis and Prevention*, 42(2), 492–9. doi:10.1016/j.aap.2009.09.013.
- Scott, P. P. (1986). Modelling time-series of British road accident data. *Accident Analysis and Prevention*, 18(2), 109–17.
- Scuffham, P.A., & Langley, J.D. (2002). A model of traffic crashes in New Zealand. *Accident Analysis and Prevention*, 34(2002), 673–87.

- Sen, A., & Mizzen, B. (2007). Estimating the impact of seat belt use on traffic fatalities: empirical evidence from Canada. *Canadian Public Policy / Analyse de Politiques*, 33(3), 315–35.
- Shankar, V., Mannering, F., & Barfield, W. (1996). Statistical analysis of accident severity on rural freeways. *Accident Analysis and Prevention*, 28(3), 391–401.
- Shefer, D. (1994). Congestion, air pollution, and road fatalities in urban areas. *Accident Analysis and Prevention*, 26(4), 501–9.
- Shen, Y., Hermans, E., Brijs, T., Wets, G., & Vanhoof, K. (2012). Road safety risk evaluation and target setting using data envelopment analysis and its extensions. *Accident Analysis and Prevention*, 48, 430–41. doi:10.1016/j.aap.2012.02.020.
- Simonet, S., & Wilde, G. J. S. (1997). Risk: perception, acceptance and homeostasis. *Applied Psychology*, 46(3), 235–252. doi:10.1111/j.1464-0597.1997.tb01228.x.
- Singleton, W.T. (1984). Future trends in accident research in European countries. *Journal of Occupational Accidents*, 6, 3–12.
- Smeed, R. J. (1949). Some statistical aspects of road safety research. *Journal of the Royal Statistical Society. Series A (General)*, 112(1), 1–34.
- Smeed, R. J. (1952). Some factors influencing the road behaviour of vehicle drivers. *Operational Research Quarterly*, 3(4), 60–7.
- Smeed, R. J. (1961). The influence of speed and speed regulations on traffic flow and accidents. *International Road Safety and Traffic Review*, 9(1), 51–66.
- Soole, D., Watson, B., & Alexia, L. (2009). The impact of police speed enforcement practices on self-reported speeding: An exploration of the effects of visibility and mobility. In: *Proceedings of the 2009 Australasian Road Safety Research, Policing and Education Conference: Smarter, Safer Directions*, 10–12 November 2009. Sydney.
- Stamatiadis, N., & Deacon, J. A. (1995). Trends in highway safety: effects of an aging population on accident propensity. *Accident Analysis and Prevention*, 27(4), 443–59.
- Stipdonk, H., Bijleveld, F., van Norden, Y., & Commandeur, J. (2013). Analysing the development of road safety using demographic data. *Accident Analysis and Prevention*, 60, 435–44. doi:10.1016/j.aap.2012.08.005.
- Stipdonk, H., Wesemann, P., & Ale, B. (2010). The expected number of road traffic casualties using stratified data. *Safety Science*, 48(9), 1123–33. doi:10.1016/j.ssci.2010.04.010.

- Stolzenberg, L., & D'Alessio, S. J. (2003). "Born to be wild": the effect of the repeal of Florida's mandatory motorcycle helmet-use law on serious injury and fatality rates. *Evaluation Review*, 27(2), 131–50. doi:10.1177/0193841X02250524
- Stutts, J. C., & Martell, C. (1992). Older driver population and crash involvement trends, 1974-1988. *Accident Analysis and Prevention*, 24(4), 317–27.
- Sukhai, A., Jones, A. P., Love, B. S., & Haynes, R. (2011). Temporal variations in road traffic fatalities in South Africa. *Accident; Analysis and Prevention*, 43(1), 421–8. doi:10.1016/j.aap.2010.09.012.
- Supramanian, V. (1984). Fatal motorcycle accidents and helmet laws in Peninsular Malaysia. *Accident Analysis and Prevention*, 16(3), 157–62. doi:10.1016/0001-4575(84)90009-5.
- Tihanskv, D. P. (1974). Impact of the on traffic energy crisis accidents. *Transpn Res.*, 8, 481–92.
- Tourin, B. (1958). Ejection and automobile fatalities. *Public Health Rep.*, 73(5): 381–91.
- Travis, L. L., Clark, D. E., Haskins, A. E., & Kilch, J. A. (2012). Mortality in rural locations after severe injuries from motor vehicle crashes. *Journal of Safety Research*, 43(5-6), 375–80. doi: 10.1016/j.jsr.2012.10.004.
- Traynor, T. L. (2008). Regional economic conditions and crash fatality rates—a cross-county analysis. *Journal of Safety Research*, 39(1), 33–9. doi:10.1016/j.jsr.2007.10.008.
- Trichopollos. D., Tsachageas. A., Papadakis, J., & Kalapothaki, V. (1975). Factors related to mortality from motor vehicle accidents in European countries in 1970. *Accident Analysis and Prevention*, 7, 9–13.
- Tsai, Y. J., Wang, J. D., & Huang, W. F. (1995). Case-control study of the effectiveness of different types of helmets for the prevention of head injuries among motorcycle riders in Taipei, Taiwan. *American Journal of Epidemiology*, 142(9), 974–81.
- Tsui, C. K., Rice, T. M., & Pande, S. (2014). Predictors of nonstandard helmet use among San Francisco Bay-area motorcyclists. *Traffic Injury Prevention*, 15(2), 151–5. doi:10.1080/15389588.2013.796041.
- Turner, A., & Havey, J. (1953). *Helmet*. US Patent No. 2,634,415. Washington, DC: U.S. Patent and Trademark Office.
- UN ESCAP (2011). Status paper on road safety in Malaysia. *UN ESCAP Expert Group Meeting*. Bangkok.

- Van Beeck E. F., Mackenbach J. P., Looman C. W., Kunst A. E. (1991). Determinants of traffic accident mortality in The Netherlands: a geographical analysis. *Int J Epidemiol*, 20(3), 698–706.
- Vernon, D. D., Cook, L. J., Peterson, K. J., & Michael Dean, J. (2004). Effect of repeal of the national maximum speed limit law on occurrence of crashes, injury crashes, and fatal crashes on Utah highways. *Accident Analysis and Prevention*, 36(2), 223–9. doi:10.1016/S0001-4575(02)00151-3.
- Votey, H. L. (1984). Recent evidence from Scandinavia on deterring alcohol impaired driving. *Accident Analysis and Prevention*, 16(2), 123–38.
- Votey, H. L., (1986). Taking account of system interactions in modelling road accidents. *Accident Analysis and Prevention*, 18(2), 85–94.
- Wagenaar, a C., Streff, F. M., & Schultz, R. H. (1990). Effects of the 65 mph speed limit on injury morbidity and mortality. *Accident Analysis and Prevention*, 22(6), 571–85.
- Wagenaar, A. C. (1984). Effects of macroeconomic conditions on the incidence of motor vehicle accidents. *Accident Analysis & Prevention*, 16(3), 191–205. doi:10.1016/0001-4575(84)90013-7.
- Wagenaar, A. C., & Maybee, R. G. (1986). The legal minimum drinking age in Texas: Effects of an increase from 18 to 19. *Journal of Safety Research*, 17(4), 165–78. doi:10.1016/0022-4375(86)90067-8.
- Wagenaar, A. C., Maldonado-Molina, M. M., Ma, L., Tobler, A. L., & Komro, K. A. (2007). Effects of legal BAC limits on fatal crash involvement: analyses of 28 states from 1976 through 2002. *Journal of Safety Research*, 38(5), 493–9. doi:10.1016/j.jsr.2007.06.001.
- Wagenaar, A. C., Maybee, R. G., & Sullivan, K. P. (1988). Mandatory seat belt laws in eight states: a time-series evaluation. *Journal of Safety Research*, 19(2), 51–70. doi:10.1016/0022-4375(88)90044-8.
- Wang, C., Quddus, M. A., & Ison, S. G. (2013). The effect of traffic and road characteristics on road safety: A review and future research direction. *Safety Science*, 57, 264–75. doi:10.1016/j.ssci.2013.02.012.
- Webster F. V., & Newby R. F. (1964). Research into relative merits of roundabouts and traffic controlled intersections. *Proc. Inst. Civ. Engrs.*, 27, 47–75.
- Wegman, F., & Oppe, S. (2010). Benchmarking road safety performances of countries. *Safety Science*, 48(9), 1203–11. doi:10.1016/j.ssci.2010.02.003.

- Wegman, F., Aarts, L., & Bax, C. (2008). Advancing sustainable safety. *Safety Science*, 46(2), 323–43. doi:10.1016/j.ssci.2007.06.013.
- Wei, W. (2006). *Time series analysis* (2nd Ed.). Pearson Education Inc.
- Weijermars, W., & Wesemann, P. (2013). Road safety forecasting and ex-ante evaluation of policy in the Netherlands. *Transportation Research Part A: Policy and Practice*, 52, 64–72. doi:10.1016/j.tra.2013.06.001.
- WHO (2004). *World report on road traffic injury prevention*. Geneva: World Health Organization.
- WHO (2008). *World health statistics 2008*. Geneva: World Health Organization.
- WHO (2009). *Global status report on road safety: Time for action*. Geneva: World Health Organization.
- WHO (2011). *Global burden of disease*. Geneva: World Health Organization.
- WHO (2013). *Global status report on road safety 2013: supporting a decade of action*. Geneva: World Health Organization.
- Wilde, G. J. S. (1982). The theory of risk homeostasis: Implications for safety and health. *Risk Analysis*, 2(4), 209–25. doi:10.1111/j.1539-6924.1982.tb01384.x.
- Williams, A. F., Reinfurt, D., & Wells, J. K. (1996). Increasing seat belt use in North Carolina. *Journal of Safety Research*, 27(1), 33–41. doi:10.1016/0022-4375(95)00027-5.
- Wilmot, C. G., & Khanal, M. (1999). Effect of speed limits on speed and safety: a review journal. *Transport Reviews*, 19(4), 315–29. doi:10.1080/014416499295420
- Wilson, D. C. (1989). *The effectiveness of motorcycle helmets in preventing fatalities*. Washington DC: National Highway Traffic Safety Administration.
- Wintemute, G. J. (1985). Is motor vehicle-related mortality a disease of development? *Accident; Analysis and Prevention*, 17(3), 223–37.
- Wittenberg, P., Sever, K., Knoth, S., Sahin, N., & Bondarenko, J. (2013). Determining quantitative road safety targets by applying statistical prediction techniques and a multi-stage adjustment procedure. *Accident Analysis and Prevention*, 50, 566–77. doi:10.1016/j.aap.2012.06.006.
- Wong, S. C., Sze, N. N., Yip, H. F., Loo, B. P. Y., Hung, W. T., & Lo, H. K. (2006). Association between setting quantified road safety targets and road fatality reduction, *Accident Analysis and Prevention*, 38, 997–1005. doi:10.1016/j.aap.2006.04.003.

- World Bank, (2014). *World databank*. Retrieved from <http://databank.worldbank.org/ddp/home.do>.
- Xuequn, Y., Ke, L., Ivers, R., Wei Du, W., & Senserrick, T. (2011). Prevalence rates of helmet use among motorcycle riders in a developed region in China. *Accident Analysis and Prevention*, 43, 214–9. doi:10.1016/j.aap.2010.08.012
- Yaacob, W.F.W., Husin, W.Z.W., Abd. Aziz, N., & Nordin, N.I. (2011). An intervention model of road accidents: the case of Ops Sikap intervention. *Journal of Applied Sciences*, 11(7), 1105–12. doi:10.3923/jas.2011.1105.1112
- Yaffee, R. A., & McGee, M. (2000). *Introduction to time series analysis and forecasting: with Applications of SAS and SPSS*. Academic Press Inc.
- Yan-hong, L., Rahim, Y., Wei, L., Gui-xiang, S., Yu, Y., & Zhou, D. D. (2006). Field data: a study on trend and prediction of fatal traffic injuries prevalence in Shanghai. *Traffic Injury Prevention*, 7, 403–7. doi:10.1080/15389580600943336.
- Yannis, B. G., & Karlaftis, M. G. (2009). *Weather effects on daily traffic accidents and fatalities: a time series count data approach*. Athens, Greece: National Technical University of Athens.
- Yannis, G., Antoniou, C., & Evgenikos, P. (2010). Comparative analysis of junction safety in Europe, in *Proc of the 12th WCTR*. July 11–15, 2010 Lisbon, Portugal.
- Yannis, G., Antoniou, C., & Papadimitriou, E. (2007). Road casualties and enforcement: distributional assumptions of serially correlated count data. *Traffic Injury Prevention*, 8(3), 300–8. doi:10.1080/15389580701369241.
- Yannis, G., Antoniou, C., & Papadimitriou, E. (2011a). Autoregressive nonlinear time-series modeling of traffic fatalities in Europe. *European Transport Research Review*, 3(3), 113–27. doi:10.1007/s12544-011-0055-4.
- Yannis, G., Antoniou, C., Papadimitriou, E., & Katsochis, D. (2011b). When may road fatalities start to decrease? *Journal of Safety Research*, 42(1), 17–25. doi:10.1016/j.jsr.2010.11.003.
- Yannis, G., Papadimitriou, E., & Folla, K. (2014). Effect of GDP changes on road traffic fatalities. *Safety Science*, 63, 42–9. doi:10.1016/j.ssci.2013.10.017.
- Zador, P. (1983). How effective; are motorcycle daytime headlight laws? *Am J Public Health*, 73(7), 808.
- Zador, P. (1985). Motorcycle headlight-use laws and fatal motorcycle crashes in the US, 1975–83. *American Journal of Public Health*, 75(5), 543–6.

- Zhang, X., Xiang, H., Jing, R., & Tu, Z. (2011). Road traffic injuries in the People's Republic of China, 1951-2008. *Traffic Injury Prevention*, 12(6), 614–20. doi:10.1080/15389588.2011.609925.
- Zhu, X. (2010). Application of composite grey BP neural network forecasting model to motor vehicle fatality risk. *2010 Second International Conference on Computer Modeling and Simulation*, 5, 236–40. doi:10.1109/ICCMS.2010.257.
- Zivković, V., Nikolić, S., Lukić, V., Zivadinović, N., & Babić, D. (2013). The effects of a new traffic safety law in the Republic of Serbia on driving under the influence of alcohol. *Accident Analysis and Prevention*, 53, 161–5. doi:10.1016/j.aap.2013.01.012.
- Zlatoper, T. J. (1989). Models explaining motor vehicle death rates in the United States. *Accident Analysis and Prevention*, 21(2), 125–54.